AMERICAN JOURNAL *of* PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 16, NUMBER 9 DECEMBER, 1948

The Search for the Neutrino through Nuclear Recoil Experiments

JAMES S. ALLEN
University of Illinois, Urbana, Illinois

IN recent years experimental evidence resulting from investigations of beta-decay has indicated that at least some of the mystery surrounding this process can be removed by the neutrino hypothesis introduced by Pauli in 1933. Probably the most obvious reason for the introduction of the neutrino is the need for the conservation of momentum and energy in beta-decay. The existence of a continuous spectrum of beta-energies from a typical radioactive element such as P^{32} appears to contradict the assumption of transitions between discrete energy levels in both the initial and final elements. According to the neutrino hypothesis, the portion of the energy and corresponding momentum not carried off by the beta-particle is removed by the neutrino. In order to be practically nondetectable this new particle must be endowed with zero charge and with a rest mass less than that of an electron.

According to present theory, negative beta-disintegration consists of the conversion of a neutron into a proton and an electron plus a neutrino. Since the neutron, proton and electron are assumed to have spins of $\frac{1}{2}$ and to obey Fermi statistics, it is necessary to assume that the neutrino also has a spin of $\frac{1}{2}$ and Fermi statistics.

Bethe[1] and Crane[2] have shown that an estimate can be made of the mass of the neutrino from a closed cycle in which a $p-n$ reaction is followed by positron emission. Haxby, Shoupp, Stephens and Wells[3] have measured the energy threshold of the reaction

$$C^{13}+p\rightarrow N^{13}+n,$$
$$N^{13}\rightarrow C^{13}+e^{+}+\nu,$$

where ν represents the neutrino. In this closed cycle the masses of C^{13} and N^{13} cancel out, and the energy available for the positron emission may be found if the energy of the bombarding proton at threshold, the neutron-proton mass difference, the mass of the positron and the mass of the electron are known. The value found for N^{13} is 1.20 ± 0.04 Mev. When this is compared with the value of 1.198 ± 0.006 Mev obtained by Lyman,[4] the energy equivalent of the mass of the neutrino is found to be very small, if not zero. Hughes and Eggler[5] have measured the energy release for the reactions $N^{14}(n, p)C^{14}$ and $He^{3}(n, p)H^{3}$. The mass of the neutrino could be estimated when the neutron-proton mass difference and the mass of the electron were substituted in the closed cycle reaction which included the mass equivalent of the energy available for the beta-decay of C^{14} or H^{3}. The neutrino mass was found to be 1 ± 25 kev from the first reaction and 4 ± 25 kev from the second.

[1] H. A. Bethe, *Elementary nuclear theory* (Wiley, 1947) Ch. VI.
[2] H. R. Crane, *Rev. Mod. Physics* **20**, 278 (1948).
[3] R. O. Haxby, W. E. Shoupp, W. E. Stephens and W. H. Wells, *Physical Rev.* **58**, 1035 (1940).
[4] E. M. Lyman, *Physical Rev.* **55**, 234 (1939).
[5] D. J. Hughes and C. Eggler, *Physical Rev.* **73**, 1242 (1948); *Physical Rev.* **73**, 809 (1948).

Konopinski[6] has noted that the beta-ray spectrum of H^3 has such a low energy that the calculated lifetime should be extremely sensitive to the neutrino rest mass. He finds that a mass of 1/30 to 1/45 of the electron mass brings the calculated lifetime into agreement with the measured value, whereas a zero mass leads to a discrepancy of a factor 10. Pruett[7] also has considered the anomalously short life (10–20 yr) for the energy release (11–15 kev) of H^3. For a half-life of 20 yr he found the neutrino rest mass to be $0.277m$ if the energy release is 11 kev and $0.054m$ for a 15 kev energy release, where m represents the rest mass of the electron. The neutrino mass would be zero if the energy release were as high as 20 kev.

Cook, Langer and Price[8] have obtained information regarding the mass of the neutrino from an analysis of the beta-spectrum of S^{35}. This analysis is based upon the fact that the shape of the Fermi plot of beta-ray spectrum near the high energy end is sensitive to the mass of the neutrino. The reasons for this correlation have been discussed by Bethe[9] and more recently by Kofoed-Hansen.[10] In the case of S^{35} the Fermi plot is a straight line for all beta-ray energies above 60 kev. From a comparison of the observed data near the end point of the spectrum with the theoretical Fermi curves for an allowed transition calculated for assumed neutrino masses of zero, one percent and two percent of the mass of the electron, the authors concluded that the neutrino mass is less than one percent of that of the electron. About all that can be concluded from the experiments discussed in this section is that the neutrino mass is considerably smaller than that of the electron.

The most fruitful method of investigating the properties of the neutrino is a rather indirect one. In this indirect type of experiment it is assumed that the neutrino transfers a measurable momentum to the parent nucleus during the emission process. In practice the energy and momentum of the recoiling nucleus and of all other particles emitted during the decay process, with the exception of the neutrino, can be measured. In general it is found that energy and momentum are not conserved unless the neutrino is introduced to carry away the missing momentum and energy. In the remaining sections of this review the results of recent recoil experiments will be presented and analyzed for possible evidence of the neutrino. The results of earlier neutrino experiments have been discussed by Crane[2] and Sherwin.[11]

Nuclear Recoils from the *K*-Capture Process

According to our present picture of the K-electron capture process an electron from the K shell, or less frequently from an outer shell, is captured by the nucleus according to the scheme:

$$Z^A+e_k\rightarrow(Z-1)^A+\nu,$$

where Z^A represents a nucleus of atomic number Z and atomic mass A. The energy available for electron capture is

$$E=M_a(Z^A)-M_a(Z-1)^A,$$

where the phrase "*the mass, in energy units, of*" is symbolized by M_a. Whenever $E>0$ in the case of the light elements for which the binding energies of the extra nuclear electrons are small compared to nuclear binding energies, this excess energy and corresponding momentum presumably is carried off by the neutrino. As pointed out by Crane,[2] the measurement of the recoil in a K-capture process is the one experiment which can distinguish sharply between the emission of single and multiple neutrinos. If the single neutrino picture is correct, the momentum spectrum of recoils will be a line spectrum since the energy of the transformation is not shared between an electron and a neutrino, but is taken by the neutrino alone. If no gamma-rays are present, the recoil spectrum will consist of a single line. In contrast, the multiple neutrino picture would give a continuous distribution of recoil momenta.

Kan Chang Wang[12] first suggested the use of Be^7 for a recoil experiment, and an experiment was completed several months later by Allen.[13] Be^7 has been investigated by Rumbaugh, Roberts

[6] E. J. Konopinski, *Physical Rev.* **72**, 518 (1947).
[7] J. R. Pruett, *Physical Rev.* **73**, 1219 (1948).
[8] C. Sharp Cook, L. M. Langer and H. C. Price, Jr., *Physical Rev.* **73**, 1395 (1948).
[9] H. A. Bethe and R. F. Bacher, *Rev. Mod. Physics* **8**, 82 (1936), Sec. 40.
[10] O. Kofoed-Hansen, *Physical Rev.* **71**, 451 (1948).
[11] C. W. Sherwin, *Nucleonics*, May, 1948, p. 16.
[12] Kan Chang Wang, *Physical Rev.* **61**, 97 (1942).
[13] J. S. Allen, *Physical Rev.* **61**, 692 (1942).

and Hafstad.[14] They found that this isotope is formed according to the reaction

$$Li^6+D^2 \rightarrow Be^7+n \ (3.3 \text{ Mev}),$$

and that it decays in two ways, 90 percent according to

$$Be^7+e_k \rightarrow Li^7+\nu \ (0.87 \text{ Mev}),$$

and 10 percent according to

$$Be^7+e_k \rightarrow Li^{7*}+\nu \ (0.396 \text{ Mev}),$$
$$Li^{7*} \rightarrow Li^7+\gamma \ (0.474 \text{ Mev}).$$

The energy available for the neutrino and gamma-ray emission has been measured by Haxby, Shoupp, Stephens and Wells[15] and by Zaffarano, Kern and Mitchell.[16]

It turns out that in an ideal recoil experiment 90 percent of the Li^7 recoils would have a sharply defined energy of 58 ev. The remaining 10 percent, emitted in coincidence with the 0.474-Mev gamma-ray, should have a continuous distribution of energies from 58 ev to nearly zero. In the actual Be^7 experiment the active material was deposited by means of a selective evaporation process upon a strip of platinum. Since the work function of platinum is considerably larger than the first ionization potential of lithium, most of the recoils were ionized upon leaving the surface. The detecting equipment used in this experiment consisted of an electron multiplier tube for counting the recoil ions and a Geiger counter for gamma-ray counting. The maximum energy of the recoils was measured by a retarding potential method.

The experimental curves obtained by this method are shown in Fig. 1. The actual values of the upper limits of the recoil curves are 10 to 15 volts lower than the expected value. A large part of this discrepancy may be accounted for by the work required to remove the recoil ions from the platinum strip. The fact that recoils were observed with approximately the correct energy is rather strong confirmation of the neutrino hypothesis. Since the Be^7 was most certainly a "thick" source, the shape of the retarding potential curves has little connection with the true recoil energy distribution. Hence, as noted by Crane,[2] this experiment cannot differentiate between the existence of single or multiple neutrinos. A separate experiment in which the number of gamma-ray counts in coincidence with recoil counts was measured indicated that few of the observed recoils were due to gamma-ray emission alone.

One other K-capture experiment has been reported. This experiment, using 6.7-hr Cd^{107}, was initiated by Alvarez, Helmholz and Wright[17] and was contiuued by Wright.[18]

The decay scheme has been given by Bradt[19] *et al.* and is reproduced in Fig. 2. According to this scheme K-electron capture to a metastable level in Ag^{107} occurs in 99.27 percent of the disintegrations. A silver recoil ion of 7.9-ev energy is expected from this K-capture process. In Wright's experiment a Cd^{107} surface was prepared by a double vacuum distillation. This active surface was exposed to a collector for a short time

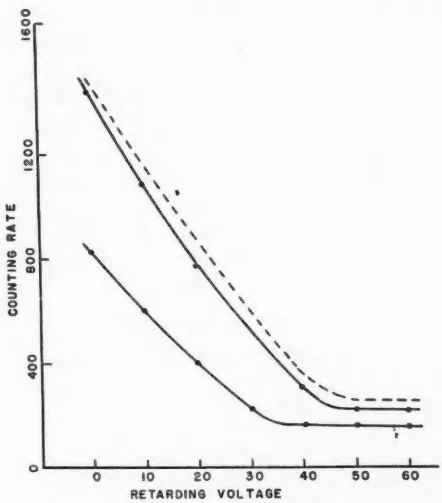

FIG. 1. Retarding potential curves for recoil ions produced by K-capture in Be^7. The upper solid curve represents data taken one hour after the Be^7 source had been outgassed. The lower curve represents data taken two hours after the heating of the source. The dotted curve represents the data corrected back to zero age for the source.

[14] L. H. Rumbaugh, R. B. Roberts and L. R. Hafstad, *Physical Rev.* **54**, 657 (1938).
[15] R. O. Haxby, W. E. Shoupp, W. E. Stephens and W. H. Wells, *Physical Rev.* **58**, 1035 (1940).
[16] D. J. Zaffarano, B. D. Kern, and A. N. Mitchell, *Physical Rev.* **74**, 105 (1948).
[17] L. W. Alvarez, A. C. Helmholz and B. T. Wright, *Physical Rev.* **60**, 160 (1941).
[18] B. T. Wright, *Physical Rev.* **71**, 839 (1947).
[19] H. Bradt, P. C. Gugelot, O. Huber, H. Medicus, P. Preiswerk, P. Scherrer, and R. Steffen, *Helv. Physica Acta* **19**, 218 (1946).

TABLE I.

Interaction	Correlation function
Scalar	$P_1(\theta)=1-(v/c)\cos\theta$
Polar vector: Fermi	$P_2(\theta)=1+(v/c)\cos\theta$
Tensor: Gamow, Teller	$P_3(\theta)=1+(v/3c)\cos\theta$
Axial vector: Gamow, Teller	$P_4(\theta)=1-(v/3c)\cos\theta$
Pseudo scalar	$P_5(\theta)=1-(v/c)\cos\theta$.

in order to collect any Ag^{107} atoms which recoiled from the Cd surface. Detection of the Ag^{107} atoms was accomplished by counting the conversion electrons. By a comparison of the amount of 6.7-hr Cd activity on the source with the activity of the Ag^{107} on the collector an efficiency of 8 percent for the collection of the recoils was indicated. That is, of all the Ag^{107} nuclei recoiling into the hemisphere available for collection, more than 8 percent were actually collected. The fact that the collection efficiency was independent of the sign of the electric field between the source and collector indicated that the recoils were electrically neutral on leaving the surface of the source. Since the first ionization potential of Ag is 2 to 3 volts higher than the work function of the clean tungsten surface used as a backing, few, if any, of the Ag recoils would be ionized on leaving the surface of the source. This is in agreement with the results obtained by Sherwin which will be discussed later in this review. In his experiments the ratio of charged to neutral recoils was strongly influenced by the nature of the substrate upon which the active material was deposited.

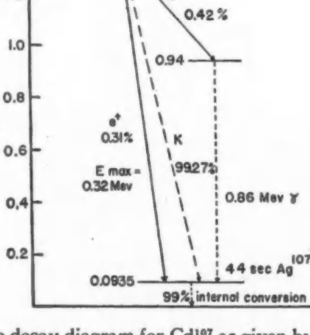

FIG. 2. The decay diagram for Cd^{107} as given by Bradt and his co-workers.

The 8 percent collection efficiency led Wright to conclude that the Ag^{107} recoils resulted from the K-capture in the 6.7-hr Cd^{107} and not from the positrons or gamma-rays produced in less than one percent of the transitions. The possibility that the recoils were directly due to the momentum imparted by the x-rays or Auger electrons following K-capture was ruled out because of the very small recoil energies expected. However, Cooper[20] has considered in detail more complicated mechanisms whereby part of the energy released during the readjustment of the electronic levels of the atom after the K-capture may be transformed into kinetic energy of the recoiling atom, the surface taking momentum. It is evident that further theoretical and experimental work is needed to determine whether or not this mechanism is possible in a K-capture experiment. In view of this uncertainty we cannot safely conclude that the observed recoils are explained by the neutrino hypothesis.

Angular Correlation Functions

Bloch and Moller[21] have noted an expected correlation between the directions of emission of the electron and neutrino in beta-decay; for allowed Fermi transitions they predict a correlation function $P(\theta)=1+(v/c)\cos\theta$ for the relative probability of decay with an angle θ between these two directions. In this expression v is the velocity of the electron and c is the velocity of light.

Hamilton[22] has extended these calculations to the allowed and first forbidden transitions corresponding to the five forms of beta-interactions which are invariant under Lorentz transformations. The correlation functions and the names that describe them for allowed transitions are given in Table I. Hamilton has emphasized that, even for allowed transitions, the correlation functions (in contrast to the energy spectra) are markedly different for the various interactions.

The recoil distributions predicted by the first two correlation functions, and also that expected in the absence of any correlation between the direction of emission of the electron and neutrino, are shown in Fig. 3. The curves have been calcu-

[20] E. P. Cooper, *Physical Rev.* **61**, 1 (1942).
[21] F. Bloch and C. Moller, *Nature* **136**, 912 (1935).
[22] D. R. Hamilton, *Physical Rev.* **71**, 456 (1947).

lated by Sussholtz[23] for a beta-emitter with $Z \leqslant 20$ and with a beta-ray spectrum of maximum kinetic energy 2 Mev. It is evident that at both the low and high energy ends of the recoil spectrum the shape of the spectrum is strongly influenced by the type of angular correlation used in the calculations. Because of this strong dependence there is hope that accurate studies of the shape of the recoil spectra from beta-emitters will yield significant information regarding the type of beta-interaction and the proper choice of selection rules.

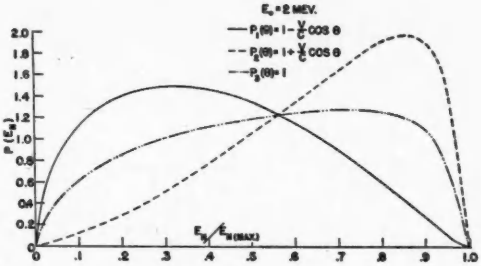

FIG. 3. Recoil distribution curves computed for three different neutrino-electron angular correlation functions. The energy, E_N, of the nuclear recoil is expressed as a fraction of the maximum recoil energy, $E_{N(\max)}$.

Nuclear Recoils from Gaseous Sources

It has been realized for a long time that a radioactive monatomic gas at a pressure of 10^{-4} mm of mercury or less would be a nearly perfect source for a nuclear recoil experiment. In this case the interpretation of the distribution of the momenta of the recoils would not be confused by the possible effects of chemical bonds in the source or by loss of energy through collisions in the gas. However, the results of investigations to be described in this section indicate that the accuracy of this type of experiment is limited by the uncertainty in the determination of solid angles due to the large volume of the source.

A study of the nuclear recoils resulting from the decay of Kr^{88} has been made by Jacobsen and Kofoed-Hansen.[24] The Kr^{88} was produced as a fission product of uranium. In order to allow the short-lived products to decay, the inert gases were not removed from the uranium until 3 hr had elapsed after the end of the neutron bombardment. The collected gas was found to consist almost entirely of Kr^{88} which decays as follows:

$$Kr^{88} \xrightarrow[2.7\text{ hr}]{2.4\text{ Mev}} Rb^{88} \xrightarrow[17.8\text{ min}]{5\text{ Mev}} Sr^{88}.$$

By means of beta-gamma-coincidence counting the presence of an isomeric transition to an excited state of Rb^{88}, followed by the emission of a gamma-ray, was established. However, neither the branching ratio nor the energy of the gamma-ray was measured.

A sketch of the apparatus used in the recoil experiment is shown in Fig. 4. The radioactive Kr^{88} gas was introduced at a pressure sufficiently low to prevent energy losses by collisions of the recoils with the gas in the chamber. Recoil Rb^{88} ions produced through the decay of Kr^{88} were collected on two foils placed opposite the open and closed sides of the inner chamber. By means of a retarding potential the recoils from the interior of the inner chamber could be prevented from reaching the foil at the right. The relative number of recoils collected on the foils was measured by means of the Rb^{88} activity. In order to correct for the recoils collected from the region between the two chambers, the difference between the two foil activities was taken to be the activity due to recoils from within the inner

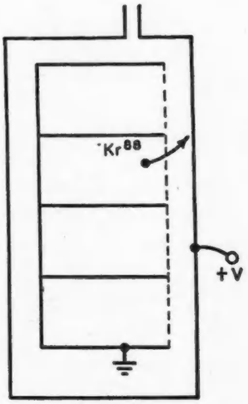

FIG. 4. A schematic diagram of the recoil apparatus used by Kofoed-Hansen. The Kr^{88} gas is contained both in the inner and outer chambers. Recoil ions are collected on foils placed on the inner surface of the outer chamber.

[23] B. Sussholtz (unpublished results).
[24] J. C. Jacobsen and O. Kofoed-Hansen, *Det. Kgl. Danske Vidensk. Selskab, Math-Fys. Med.* 23, paper 12 (1945); *Physical Rev.* **73**, 675 (1948).

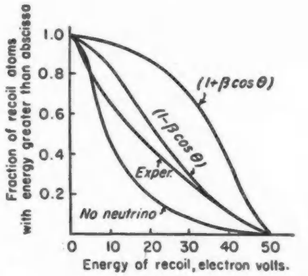

FIG. 5. The experimental and theoretical integral bias curves given by Jacobsen and Kofoed-Hansen. The theoretical curves have been computed with the assumption of a neutrino-electron angular correlation of $1+(v/c)\cos\theta$ and $1-(v/c)\cos\theta$ and also for the case of no neutrino.

chamber. The interpretation of the retarding potential curves was complicated by the existence of the excited state of Rb^{88} and also by the inherent difficulties involved in the use of the retarding potential method with an extended source. In the absence of exact information regarding the Rb^{88} level no corrections could be made for this effect. The errors introduced through the use of the retarding potential method were likewise difficult to estimate. Instead of attempting these corrections the authors show by a quantitative argument that the corrected curve must be above the uncorrected curve. Figure 5 shows the curves obtained by Jacobsen and Kofoed-Hansen. The curve labeled "no neutrino" was calculated upon the assumption that all the transitions go to the ground state of Rb^{88} and that momentum is conserved between the electron and the recoiling nucleus. The two upper curves were calculated again with the assumption that all the transitions go to the ground state of $Rb,^{88}$ but with neutrino electron angular correlation functions of $(1+\beta\cos\theta)$ and $(1-\beta\cos\theta)$. The maximum recoil energy of 51 ev is in good agreement with the value of 51.5 ev expected for the recoils from a 2.4-Mev beta-ray. This agreement does not necessarily indicate the presence of a neutrino since the energy of the neutrino, if it existed, would be zero at the high energy end of the beta-spectrum, and the contribution to the recoil momentum would also be zero. It is evident that the curve representing the observed data lies above the "no neutrino" curve and that any corrections would raise the observed data even higher. Jacobsen and Kofoed-Hansen have concluded that the experimental curve is incompatible with the assumption that there is no neutrino and that it is not possible to say anything about the angular distribution of the emitted neutrinos.

A recoil experiment using He^6 has been carried out by the author and, since a more detailed account will be reported elsewhere, only a brief summary of the results will be presented here. He^6 can be produced through the reaction

$$Be^9+n\rightarrow He^6+He^4$$

and decays according to the scheme[25,26]

$$He^6\rightarrow Li^6+e^-+\nu \quad (3.5\ \text{Mev})$$

with a half-life of 0.8 sec. Since the change of nuclear spin during the decay is probably from $I=0$ to $I=1$, the transition is forbidden according to the selection rules of the original Fermi theory of beta-decay. In order to explain the fact

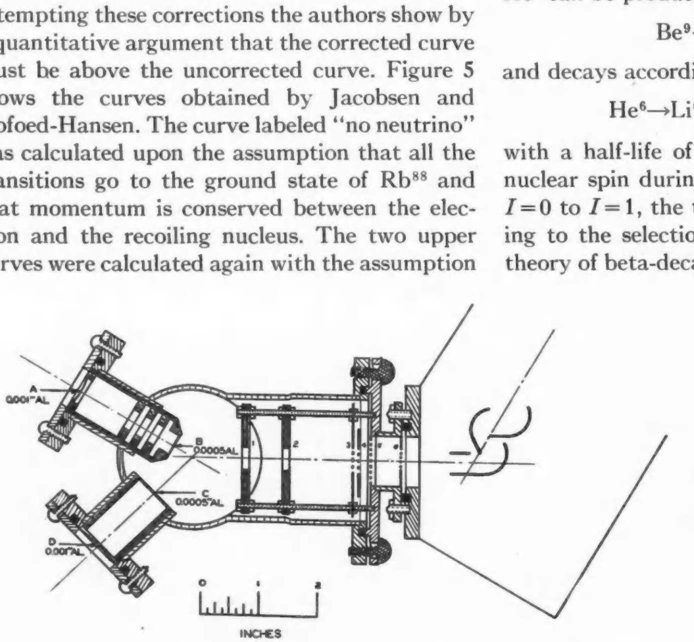

FIG. 6. A schematic diagram of the recoil apparatus used by Allen for the He^6 experiment. The beta-ray counter was placed at window A to record the data discussed in the text.

[25] T. Bjorge and K. Brostrum, *Nature* **138**, 400 (1936); *Dansk. Math. Phys. Medd.* **16**, 8 (1938).
[26] R. Sherr, *Physical Rev.* **69**, 21 (1946).

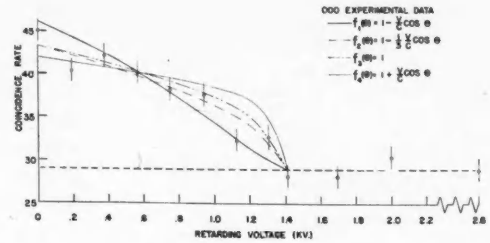

FIG. 7. Experimental and theoretical retarding voltage curves given by Allen. The beta-ray counter was in a position to record electrons emitted at $180° \pm 15°$ with respect to the recoil nuclei from the decay of He^6. The upper and lower solid curves were computed, respectively, for neutrino-electron angular correlation functions of $1+(v/c)\cos\theta$ and $1-(v/c)\cos\theta$.

that the experimental lifetime indicates an allowed transition, it is necessary to assume Gamow-Teller selection rules which permit an allowed transition for $\Delta I = 1, 0$. Since either the tensor or the axial vector form of interaction gives allowed transitions for $\Delta I = 1, 0$, the angular correlation between the directions of emission of the electron and neutrino might be expected to be of the form

$$P_3(\theta) = 1 + (v/3c)\cos\theta$$

or

$$P_4(\theta) = 1 - (v/3c)\cos\theta$$

as discussed earlier in this review.

In the actual experiment the He^6 was produced by neutron bombardment of one pound of beryllium powder in a can placed near the target of the University of Chicago cyclotron. The He^6 gas was swept out of the powder by ethyl alcohol vapor and carried through a pipe line to the recoil chamber. After the alcohol vapor was condensed in a trap cooled by liquid nitrogen, the radioactive gas diffused into the recoil chamber. The pressure in the chamber during the experiment was 10^{-5} to 10^{-6} mm of mercury. A diagram of the recoil chamber is shown in Fig. 6. An electron multiplier tube was used to count the Li^6 recoil ions, and an end-window Geiger counter was placed at either A or D to record the beta-rays. By means of a double coincidence circuit with an adjustable time delay, coincidences between the beta-rays and the corresponding recoil nuclei were recorded. The energy spectrum of the recoils was obtained by recording the coincidence rate as a function of the retarding potential between grid number 4 and the grounded grids 3 and 5. In order to indicate the amount of correction to be applied to the observed recoil spectrum because of the distortion introduced through the use of the retarding potential method, an auxiliary experiment was performed. An extended source of Li ions was placed at the intersection of the center lines through the two beta-ray ports A and D and that through the grid system. The Li ions were accelerated as a broad beam through the grid system into the multiplier tube and retarding potential curves recorded for various accelerating potentials. The family of retarding potential curves obtained in this manner was used to correct the shape of the computed curves rather than to correct the observed data.

The experimental data obtained with a recoil chamber similar to that of Fig. 6, but having a beta-ray counter in a position to record electrons emitted at $180° \pm 15°$ with respect to the recoil nuclei, are shown in Fig. 7. The curves predicted by four different angular correlations are also given. In this case the shapes of the computed curves were corrected to allow for the effect of the retarding potential on the observed data. Because of the inaccuracy of the data it is not safe to say more than that the $1+(v/c)\cos\theta$, $1+(v/3c)\cos\theta$ and probably the $1-(v/c)\cos\theta$

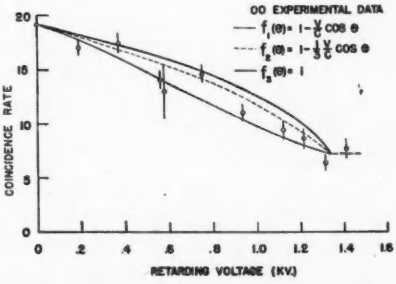

FIG. 8. Experimental and theoretical retarding voltage curves given by Allen. The beta-ray counter was in a position to record electrons emitted at an angle of $162° \pm 8°$ with respect to the direction of the nuclear recoils. The upper and lower solid curves were computed, respectively, for neutrino-electron angular correlation functions of 1 and $1-(v/c)\cos\theta$.

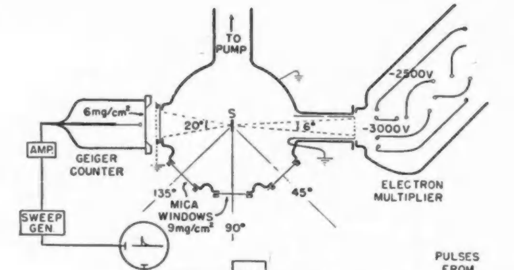

FIG. 9. A schematic diagram of the recoil apparatus used by Sherwin. The Geiger counter could be placed at either the 180°, 135°, 90° or 45° positions.

correlation are ruled out. The best agreement is with the $1-(v/3c)\cos\theta$ correlation, although the random distribution is not entirely ruled out. The retarding potential curve expected with the assumption of conservation of momentum between the electron and the recoiling nucleus in the absence of the neutrino drops rapidly towards zero as the retarding potential is increased and lies far below the experimental data at the high energy end of the recoil spectrum. Hence, it is safe to conclude that the experimental data strongly contradict the "no neutrino" assumption.

Figure 8 represents the data obtained with the arrangement shown in Fig. 6 and with the Geiger counter at position A. The data represented by open circles were recorded with the series of slits removed from the tube in front of the Geiger counter. Since additional experiments demonstrated that most of the coincidences were due to recoils originating from the region near slit number 1, the angular spread between the directions of the electrons and the recoils was taken to be $162°\pm8°$. As in the 180° arrangement, the $1+(v/c)\cos\theta$, and probably the $1-(v/c)\cos\theta$ and random distribution functions are ruled out. The best agreement is with the $1-(v/3c)\cos\theta$ correlation. The datum indicated by an open square at 580 volts was obtained with the beta-ray defining slits in place. In addition, data were recorded at retarding potentials of 0 and 1500 volts. By an adjustment of the ordinate scale the coincident rates at 0 volts and at 1500 volts, respectively, were normalized to the values given by the open circle at 0 volts and the horizontal dotted line. The curve obtained with this arrangement is expected to lie slightly below that obtained without the defining slits. Because of the large statistical error in the last set of data, it is safe to conclude that only the $1+(v/c)\cos\theta$, $1+(v/3c)\cos\theta$ and the random distribution functions are excluded.

The results of the He^6 experiment seem to rule out the "no neutrino" assumption and give some indication of agreement with the $1-(v/3c)\cos\theta$ correlation predicted by the axial vector form of interaction which follows Gamow-Teller selection rules. As mentioned earlier, the He^6 decay is expected to be governed by this type of selection rule. It is hoped that further experiments with He^6 will lead to a more critical choice of the appropriate beta-interaction or combination of interactions.

Nuclear Recoils from Thin Sources

The availability of high activity, separated, carrier free radioactive materials from the Clinton Laboratories has made possible the preparation of highly concentrated sources. By means of vacuum evaporation these sources can be deposited as extremely thin layers on a suitable backing. Under the best conditions of preparation these sources are probably monatomic layers.

Sherwin[27, 28] has studied the recoil spectra from thin sources of P^{32} and Y^{90}. A diagram of his apparatus is shown in Fig. 9. As in the experiment of Allen, the beta-rays were recorded by an end-window Geiger counter and the recoil ions by

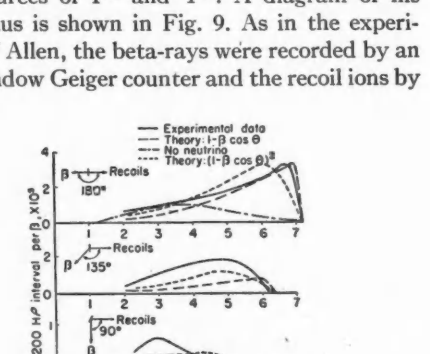

FIG. 10. Theoretical and experimental recoil momentum spectra from the decay of P^{32} as given by Sherwin.

[27] C. Sherwin, *Physical Rev.* **73**, 216 (1948).
[28] C. Sherwin, *Physical Rev.* **73**, 1173 (1948).

an electron multiplier tube. The time of flight method was used to measure the distribution in velocity of the nuclear recoils. In this method the pulse produced by the entrance of a beta-ray into the Geiger counter initiates a linear sweep on the cathode-ray tube. After an interval of time corresponding to that required for the recoil ion to travel from the source at S to the first electrode of the multiplier tube, a pip appears on the screen. From a knowledge of the sweep speed, the time of flight of the ions may be determined with considerable accuracy. The intensity of the recoils was low enough so that each recoil could be observed visually, and the time interval in which it occurred recorded.

The preparation of the thin layer, radioactive sources was the most crucial part of Sherwin's experiment. The sources were deposited by evaporation within the recoil apparatus and even with a vacuum of 10^{-6} to 10^{-7} mm of mercury, the sources deteriorated in times ranging from 20 min to 2 hr. The probable cause of this deterioration was the adsorption of residual gases and vapors upon the surface of the radioactive source.

The first experiment reported by Sherwin used P^{32} which decays according to the scheme:

$$P^{32} \rightarrow S^{32} + e^{-} + \nu \ (E_{\max} = 1.72 \text{ Mev}).$$

Since no gamma-rays are released in the decay, the recoil spectrum should be due to the momenta received from the beta-decay process. The maximum recoil energy expected is 78 electron volts. Since only charged recoil atoms could be accelerated into the multiplier tube, it was necessary to use a surface whose work function was greater than the first ionization potential of S^{32}. Since the first ionization potential of sulfur is 10.3 volts, this requirement ruled out the use of metal surfaces. However, freshly evaporated surfaces of SiO_2, LiF and NaF were found to permit some 10 percent or less of the recoil muclei to escape as ions.

Figure 10 shows the recoil curves obtained using a P^{32} source deposited on a thin layer of LiF. In addition to the experimental data the curves expected for the "no neutrino" assumption, and also the angular correlation functions $1-(v/c)\cos\theta$ and $[1-(v/c)\cos\theta]^2$, are given. It is obvious that the experimental data taken at 180° strongly contradict the "no neutrino" assumption.

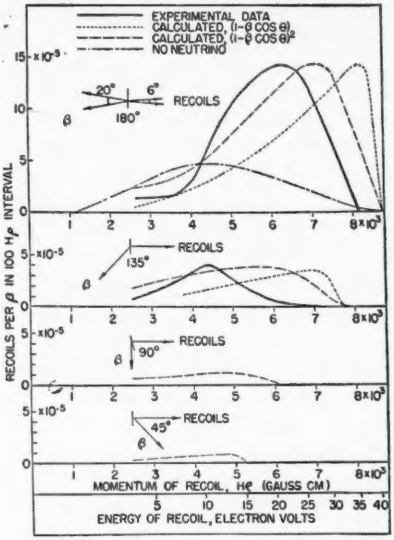

FIG. 11. A comparison of the observed recoil momentum spectra from Y^{90} with those computed from the neutrino hypothesis (after Sherwin).

The author rules out the random, $1+(v/c)\cos\theta$, and $1-(v/c)\cos\theta$ first forbidden angular correlation functions. The $1-(v/c)\cos\theta$ correlation is an excellent fit with the experimental 180° curve, but the agreement is less close as the angle decreases. In the case of the 135° curve the agreement with the $[1-(v/c)\cos\theta]^2$ correlation appears to be somewhat better than that with the $1-(v/c)\cos\theta$ correlation. The author points out that the 180° data should be given more weight than the others because the recoils have the greatest energy and because comparatively less distortion of the curve toward lower energies due to energy loss in the escape of the recoil from the surface should be expected.

Sherwin[28] has reported a second recoil experiment in which Y^{90} was used. The decay scheme is:

$$Y^{90} \rightarrow Zr^{90} + e^{-} + \nu \ (E_{\max} = 2.16 \text{ Mev})$$

with a maximum recoil energy of 40.7 electron volts. In this experiment the ionization potential of 6.92 volts for Zr was sufficiently low to allow the use of either tungsten or quartz as the backing material upon which the Y^{90} was deposited. An estimated 87 percent of the nuclear recoils were able to escape as singly charged ions. The recoil

apparatus was essentially the same as that used for the P^{32} experiment.

In Fig. 11 a comparison of the observed recoil momentum spectra from Y^{90} with those calculated from the neutrino hypothesis are shown. As in the P^{32} experiment, the curve predicted by the "no neutrino" assumption definitely does not fit the experimental data. The most noticeable discrepancy between the observed and calculated recoil spectra is that the observed spectra (for both 180° and 135°) fall short of the maximum energy expected by about 6 volts. The author suggests that a possible explanation of this energy loss is the binding energy of the Zr^{+} ions to the surface and to the oxygen atom to which they are probably attached. It is concluded that, if a 6-volt correction is made, a $[1-(v/c)\cos\theta]^2$ neutrino-electron angular correlation function accounts in an almost quantitative manner for the observed spectra. However, the $1-(v/c)\cos\theta$ correlation could not be definitely ruled out.

In regard to the important question as to the effect of the surface on the shape of the momentum spectra, Sherwin feels that no exact limits of error due to chemical bonds can be set at the present time. However, the reproducibility of the recoil spectra for different chemical combinations at the surface and the systematic disappearance of high momentum recoils at 135°, 90° and 45° suggest that the surface effects do not seriously interfere with their behavior as predicted by the neutrino hypothesis. A final selection of the neutrino-electron correlation function cannot be made until more exact measurements of surface effects have been carried out.

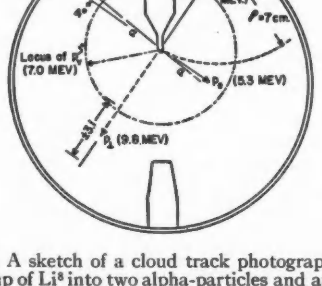

FIG. 12. A sketch of a cloud track photograph showing the breakup of Li^8 into two alpha-particles and an electron. The resultant momentum of the three particles is 9.8 ± 3.1 in Mev units where the uncertainty arises from the possibility of scattering of the alpha-particles. A neutrino with the remaining available energy of 7.0 Mev could have this momentum within the probable error (after Christy and co-workers).

Nuclear Recoils from the Decay of Li^8

Numerous investigators have proposed the study of the conservation of momentum in the decay of Li^8 which is unique in that it is followed by the breakup of the daughter nucleus, Be^8, into two alpha-particles. The decay scheme is

$$Li^8 \rightarrow Be^{8*} + e^- + \nu,$$
$$\searrow$$
$$He^4 + He^4$$

with the initial electron decay having a half-life of 0.88 sec, followed by the alpha-particle decay in a time of the order of 10^{-21} sec. The Be^8 level[29] is extremely broad, but the maximum probability is for the alpha-pair to have about 3 Mev and the electron and neutrino together to have an energy of about 12 Mev. In principle, the Li^8 may be introduced into a cloud chamber and measurements made of the energy and momenta (vector) of the two alpha-particles and of the electron. If the energy and momenta can be obtained with sufficient accuracy, the direction and momentum of the neutrino can be found by completing the momentum diagram. Although this experiment is capable of yielding rather specific information regarding the neutrino, the required accuracy has not been achieved because of the severe technical difficulties of the experiment.

Christy, Cohen, Fowler, Lauritsen and Lauritsen[30] have reported the results of a Li^8 recoil experiment which, although the results are not entirely conclusive, represents the first success with this method. The Li^8 formed through the reaction

$$Li^7 + H^2 \rightarrow Li^8 + H^1$$

was introduced on a thin foil into a cloud chamber by means of a rapidly moving plunger. A photo-

[29] W. A. Fowler and C. C. Lauritsen, *Physical Rev.* **51**, 1103 (1937). L. H. Rumbaugh, R. B. Roberts and L. R. Hafstad, *Physical Rev.* **51**, 1106 (1937); *Physical Rev.* **54**, 657 (1938). C. Smith and W. Y. Chang, *Proc. Roy. Soc.* **A166**, 415 (1938). T. Bonner, J. E. Evans, C. W. Malich and J. R. Risser, *Physical Rev.* **72**, 163A (1947).

[30] R. F. Christy, E. R. Cohen, W. A. Fowler, C. C. Lauritsen and T. Lauritsen, *Physical Rev.* **72**, 698 (1947). W. A. Fowler, C. C. Lauritsen and T. Lauritsen, *Physical Rev.* **72**, 738A (1947).

graph was taken each time the Li^8 source arrived at the center of the chamber.

A sketch of a typical cloud-chamber photograph is shown in Fig. 12. In this example the tracks of the two alpha-particles and also that of the associated electron lie in the same plane. In all, twenty-eight out of 10,000 photographs had electron tracks associated with the tracks of the two alpha-particles and could, therefore, be analyzed for "missing" momentum. The component of momentum of the two alpha-particles which was perpendicular to the average line of the two particles was measurable through the slight deviation from 180° of the angle between the two tracks. The maximum deviation from 180° was computed to be about 6°. In theory, the component of momentum parallel to the average line of the two alphas could be determined from the difference in the ranges of the two particles. However, rather large errors were introduced into the range measurements by the fact that at least one of the alphas had to pass through the LiOH target, often obliquely. After investigation of the errors to be expected, the authors concluded that only the perpendicular component of momentum could be measured with sufficient accuracy. It was possible in the 28 cases mentioned above to add the perpendicular component of the momentum of the two alphas to the corresponding momentum component of the electron and to obtain $P_\perp$, the resultant momentum of the three observable particles. This was compared with the possible total momentum of a neutrino of zero rest mass which has an energy given by the maximum energy released in the Li^8 breakup minus the observed energy of the two alphas and the electron.

An examination of the data on these 28 pictures showed that in the majority of cases the observed resultant momentum of the three observable particles was much larger than could be explained on the basis of scattering alone. In many of these cases the momentum could be understood on the neutrino hypothesis. However, the authors concluded that the data were not sufficient for a meaningful numerical comparison with any proposed assumption concerning the angular distribution of the electron and neutrino.

Conclusion

Practically all the experimental evidence indicates that there is an apparent nonconservation of momentum in the beta-decay process and that the neutrino hypothesis is at least one explanation of the missing momentum. At the present time further investigations of the recoils from the K-capture process appear to be the most feasible experiments for differentiating between the single and multiple neutrino hypothesis.

The principal field to be exploited now by means of nuclear recoil experiments is the determination of the neutrino-electron angular correlation functions. It is hoped that further studies will yield information regarding the proper choice of beta-interactions and clearer understanding of the basis for the adoption of various types of selection rules.

University of Iowa Colloquium for College Physicists

The date of the Colloquium has been set for June 16, 17 and 18, 1949, at the State University of Iowa, Iowa City, Iowa. Early announcement is made of four prizes, as follows:

For the best new *experimental* teaching devices for laboratory or classroom that are exhibited, three prizes, 25, 15, and 10 dollars. Effectiveness in teaching and in arousing student interest will take precedence over excellence of shop construction.

For the best *nonexperimental* teaching device (such as a chart, a model or a manner of presentation) that is exhibited to aid the student in the understanding of concepts and phenomena, one prize, 25 dollars.

Suggestions relating to the program, and other correspondence concerning the Colloquium will be welcomed by Professor G. W. Stewart.

A Simple Proof of Malus' Theorem in Geometrical Optics

Marcelo Alonso
Instituto del Vedado, Habana, Cuba

As is well known, Malus' theorem asserts that if a system of rays is initially orthogonal to some surface (normal rectilinear congruence), it remains orthogonal to some surface after any number of reflections or refractions. For proof it is necessary to consider only a single refraction. In the case of a reflection proof can be carried out along entirely similar lines, or can be accepted without proof by those who consider reflection as a special case of refraction from the mathematical point of view, making the index of refraction -1.

The optical length of the path of a ray which goes from point A to point B through a medium of index of refraction n, not necessarily homogeneous, is given by

$$[AB]=\int_A^B n\,dl,$$

where the integral is taken along the path of the ray. The optical length is proportional to the time taken by the light to go from A to B. If the medium is homogeneous, n is constant, the path is a straight line and l is simply the distance AB; then $[AB]=nAB$.

Consider a system of rays (Fig. 1) initially orthogonal to S, and let S'' be a reflecting or refracting surface interposed in the paths of rays, and S' another surface, locus of the end points of the rays emanating from S and having the same optical length. Then for any two adjacent rays, $[APA']=[BQB']$. Evidently, if n and n' are the indexes of refraction of the two mediums on either side of S'',

$$[APA']=nr+n'r',$$

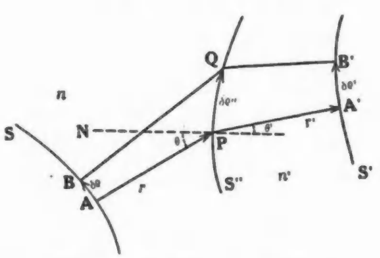

Fig. 1. Paths of rays crossing a refracting surface S''.

and since when A, P and A' are displaced to B, Q and B', respectively, the optical length remains unchanged,

$$\delta[APA']=n\delta r+n'\delta r'=0. \qquad (1)$$

But

$$r\delta r=\mathbf{r}\cdot\delta\mathbf{r}=\mathbf{r}\cdot(\delta\boldsymbol{\varrho}''-\delta\boldsymbol{\varrho})=\mathbf{r}\cdot\delta\boldsymbol{\varrho}''=r\sin\theta\delta\rho'', \qquad (2)$$

because $\mathbf{r}\cdot\delta\boldsymbol{\varrho}=0$, since AP is normal to S. Similarly,

$$r'\delta r'=\mathbf{r}'\cdot(\delta\boldsymbol{\varrho}'-\delta\boldsymbol{\varrho}'')=\mathbf{r}'\cdot\delta\boldsymbol{\varrho}'-r'\sin\theta'\delta\rho''. \qquad (3)$$

Substituting Eqs. (2) and (3) in Eq. (1),

$$(n\sin\theta-n'\sin\theta')\delta\rho''+\mathbf{r}'\cdot\delta\boldsymbol{\varrho}'/r'=0. \qquad (4)$$

The expression in parentheses is zero due to Snell's laws for refraction. Therefore,

$$\mathbf{r}'\cdot\delta\boldsymbol{\varrho}'=0,$$

showing that PA' is also normal to S'. The theorem is thus proved.

We may observe that, given a surface orthogonal to the rays, we may obtain any other surface by taking equal optical lengths along the rays from the given surface. Since equal optical lengths are equivalent to equal times of propagation, the orthogonal surfaces considered in geometrical optics correspond to the wave surfaces in physical optics. The theorem of Malus justifies the use of this concept in geometrical optics without the introduction of any special hypothesis about the structure of light.

The present proof has, in addition to its simplicity, a further advantage which shows the relation between Malus' theorem and the laws of reflection and refraction. First, we have derived it by application of these laws. Conversely, if we postulate Malus' theorem in geometrical optics we can derive from it the laws for reflection and refraction. We shall give only the proof for refraction, since the case for reflection can be proved in the same way, using the value -1 for the index of refraction.

Assuming Malus' theorem, $\mathbf{r}\cdot\delta\boldsymbol{\varrho}=0$ and

$\mathbf{r}'\cdot\delta\boldsymbol{\varrho}'=0$. Taking $\delta\boldsymbol{\varrho}''$ perpendicular to the plane made by AP and the normal PN to S'' at P, $\mathbf{r}\cdot\delta\boldsymbol{\varrho}''=0$ and Eq. (1) reduces to $\mathbf{r}'\cdot\delta\boldsymbol{\varrho}''=0$, showing that PA' also lies in the plane of AP and PN. This is the first law; namely, the incident and refracted rays and the normal to the refracting surface at the point of incidence lie in the same plane.

Taking $\delta\boldsymbol{\varrho}''$ in the plane of incidence APA', Eq. (1) reduces to (compare Eq. (4))

$$(n\sin\theta - n'\sin\theta')\delta\rho''=0.$$

Therefore,

$$n\sin\theta = n'\sin\theta',$$

which is Snell's law. (This law might better be known under the name of Descartes-Snell, paying due acknowledgment to Descartes' contribution to its mathematical formulation.) It should be noted that in many elementary books the proof of Snell's law, based ostensibly on Huygens' Principle, is formally wrong since it is really an application of Malus' theorem in the form given above.

In the development of geometrical optics there have been three different methods of approach: (1) Descartes-Snell laws, (2) Fermat's principle of stationary optical path, and (3) Malus' theorem. It is well known that if we assume either the first or the second method we can derive the third. In many books Malus' theorem is proved from Fermat's principle; the converse is also possible. In the present paper we have proved the intimate connection between the first and third methods in the sense that if we assume as true either one, the other must be necessarily true. Therefore, any one of the three statements can be considered as the fundamental postulate of geometrical optics. But since Malus' theorem is essentially a direct result of the wave theory of light, where the wave surface is the fundamental concept, and geometrical optics is merely a limiting case of physical optics, it appears to us that the most natural approach to geometrical optics is through Malus' theorem. Though this point of view does not seem practicable in elementary courses, it should be analyzed in advanced courses.

What Does Electromotive Force Mean?

JOHN A. ELDRIDGE
University of Iowa, Iowa City, Iowa

HAZELTINE[1] and Hudson[2] have discussed the relative merits of the terms *voltage* and *electromotive force*. I think we need both terms: voltage, referring to anything which might conceivably be measured with a voltmeter, and electromotive force, to which we should give a more specific meaning. It seems to me that the discussion should not be concluded without some analysis of exactly what we mean by electromotive force. Actually, the term is used in a variety of senses.

1. *Electromagnetic field theory*. The electromotive force about a closed path is defined as the line integral of the E-field around it. Expressed in symbols, the emf is $\oint E_{\|}ds$. If we adopt this definition there is an electromotive force only when there is a changing magnetic flux. A voltaic cell does not have an emf. (Note that magnetomotive force is the analog of electromotive force as defined in this sense.) When there is no emf about any circuit in the field, E may be derived from a potential and

$$\int_a^b E_{\|}ds$$

is called the difference of potential, not the emf, between the points a and b.

2. *D.C. circuits*. A source of electromotive force is an agency in which electrical potential energy in a circuit is created reversibly at the expense of

[1] A. Hazeltine, *Am. J. Physics* **15**, 191 (1947).
[2] R. G. Hudson, *Am. J. Physics* **15**, 428 (1947).

some other form of energy. Examples are static machines, voltaic cells, inductors, transformers, contact potentials, thermocouples, photronic cells. The magnitude of the electromotive force is the energy so transformed per unit charge passing through the agency; it is the difference of potential maintained across a conducting element when no current flows. If the direction of the current is reversed, the agency becomes a counter electromotive force in which electrical potential energy is transformed reversibly into mechanical energy, chemical energy, heat energy or some other form of energy. A resistor does not constitute a counter electromotive force because the heat generation is not reversible. In an inductor, magnetic energy (electric kinetic energy) is converted into electrical potential energy; thus it is included within our definition as a source of electromotive force.

In applying Kirchhoff's law for direct currents the term electromotive force is used in this sense. In a conductive circuit, operating under steady conditions the increase in potential energy of the charge as it passes through the forward electromotive forces is equal to the reversible potential energy loss in counter electromotive forces plus the irreversible loss in resistors. In a broken circuit, for example a circuit which includes a condenser, there is usually no current and the sum of the electromotive forces, including contact emf, then gives the potential difference across the break.

3. *A.C. circuits*. There is an emf in an operated static machine; the electrodes themselves are a source of electric field but not a source of electromotive force. If these electrodes are brought close together we have a condenser. There is no emf, in the sense in which we have been using the term, in a condenser. However in the theory of a.c. circuits, capacitors and inductors are simply reactors of opposite signs. So, as we analyze the circuit, if we call the potential difference across an ideal inductor an emf, we can hardly avoid using the same term for the potential difference across a capacitor. This is actually the usual practice. In a well-known textbook[3] we read that "the condenser produces an electromotive force in the circuit." Kirchhoff's Law then becomes: in any circuit (including condensers) at any instant the sum of the electromotive forces is equal to the sum of the *ir* drops.

4. *A.C. circuits, extreme usage*. Usage often goes further than this. The sinusoidal potential difference measured across any part of an a.c. circuit is likely to be called an electromotive force. For example in discussing an inductor and resistor in series the same book[4] says: "The three electromotive forces are related as the three sides of a right-angled triangle." Here the "electromotive forces" refer to the root-mean-squared values of the potential differences across the inductor and the resistance and across the combination. The *ir* is now classed as an electromotive force. If we use the term in this sense Kirchhoff's law becomes: in any a.c. circuit the sum of the electromotive forces equals zero. Since the term emf has by now lost any specific meaning which it previously had, many prefer to substitute the term voltage. We may not care for the particular term voltage but I think some rather specific meaning should be retained for the term electromotive force.

---

[3] S. G. Starling, *Electricity and magnetism* (Longmans Green, 1925), p. 310.
[4] S. G. Starling, *Electricity and magnetism* (Longmans Green, 1925), p. 356.

Lines inspired by the article, **Modern Terminology for Physics,** by P. MOON AND D. E. SPENCER, *Am. J. Physics* **16**, 100 (1948).

Through the Phantos Glass
'Twas thalpance and the kratosage
Did zenos in the kampulos:
All radiant the heliosent,
And the zenos phosage.

IRA FREEMAN (with apologies to LEWIS CARROLL).

Upper Atmosphere Temperatures from Remote Sound Measurements

Everett F. Cox*
Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

Much of today's knowledge of upper atmosphere temperatures dates from a discovery made in 1901 during the funeral of Queen Victoria. Minute-guns were fired in London as part of her funeral rites. The easily identified noise of the rhythmic gunfire was heard distinctly in London, and recognized far to the north, but the two areas of audibility were separated by a skip zone of silence.[1]

If one assumes that over England on the date of Queen Victoria's funeral a high wind gradient to the north existed at high altitudes, the observed phenomenon is conveniently explained. The speed of sound in still air depends on air temperature, which drops at a nearly constant rate of 5° to 6°C per kilometer increase in altitude through the troposphere. With the upper parts of sound wave fronts traveling at slower speed than the lower parts, sound rays bend upwards. Audible sounds do not, therefore, travel to great distances by paths near the earth unless wind gradients or temperature inversion layers hold them down. However, sound rays which have curved upward through the troposphere could be refracted earthward if the horizontal component of wind velocity in the direction of the sound rays increased with altitude. An area of audibility could thus be formed at a great distance from the source, and the near and far areas of audibility could be separated by a skip zone.

Yet, less than three years later doubt was cast upon the simple gradient theory. On December 14, 1903, the distant area of audibility associated with an explosion in Förde (in Westfalen, Germany) subtended a horizontal angle at the source of about 100°. An explosion at Witten-Annen, November 28, 1906, produced an outer zone subtending 150°. Finally, the outer audibility zone associated with an explosion in Moscow, May 9, 1920, completely encircled the source.[2] Such lack of directionality, well illustrated by Fig. 1, is not handled effectively by a theory based on vector wind gradients.

G. von dem Borne therefore offered a qualitatively reasonable, scalar explanation for the outer zone or ring of audibility.[3] As every beginning student soon learns,[4] sound travels faster in lighter gases than in air, whose molecular weight is 29.1 g/mole. If the upper atmosphere contained a high percentage of hydrogen or helium, then sound rays which had curved upward through the troposphere would be refracted downward as the lighter gas concentration increased. Sound waves would strike the earth at great distances in *all* directions from the source when no winds were present.

But as improved communication facilities, acoustic instrumentation, and planned explosions permitted accurate determinations of sound-signal traveltimes to distant points, the von dem Borne theory lost favor. Balloon flights had found the negative thermal gradient characterizing the troposphere stopped at 12- to 16-km altitude. Above this imaginary surface, named the tropopause, air temperature remained essentially constant to the summit of ballon flights. Calculations based on sound-signal traveltimes and balloon-measured atmosphere temperatures showed that the sound waves observed outside the zone of silence must have reached a height of about 40 km at the midpoint of their travel. Samples of air from the −50°C isothermal layer, also called the stratosphere, showed no significant increase in lightweight gas concentration. Indeed, modern spectrograms of the aurora borealis show very little hydrogen or helium present at *any* altitude of the upper atmosphere.[5]

Traveltimes of sound waves from the source to observers in distant areas of audibility are appreciably greater than computed traveltimes,

---

* Now with the Sandia Branch of the Los Alamos Scientific Laboratory, Albuquerque, New Mexico.

[1] Charles Davison, *Quart. Rev.* **452**, 51 (1917).

[2] The most complete account of outer audibility zones is given by A. Wegener, *Zeits. f. Geophysik* **1**, 297 (1925).

[3] G. von dem Borne, *Physikalische Zeits.* **11**, 483 (1910).

[4] R. M. Sutton, *Demonstration experiments in physics* (McGraw-Hill, 1938), p. 171.

[5] J. Kaplan, *Nature* **136**, 549 (1935); C. W. Gartlein, *Bull. Am. Physical Soc.* **23**, 11 (1948).

if ground level air temperature is employed in the Laplace equation:

$$c=(\gamma P/\rho)^{\frac{1}{2}}=(\gamma RT/M)^{\frac{1}{2}}. \tag{1}$$

Here c is the group speed of sound waves, R is the gas constant, γ is the ratio of specific heats of the gas, M is its molecular weight, and P, ρ, and T are, respectively, ambient gas pressure, density, and absolute temperature. Since apparent speeds of signals to points beyond the zone of silence are some 10 to 20 percent less than normally expected, the expressions *abnormal audibility* and *abnormal sounds* have come into common usage as descriptive of compressional waves received in the distant zones of audibility.

Today's most generally accepted theory of abnormal sound was introduced by F. J. W. Whipple[6] in 1923, but some strong skepticism yet remains.[7] Studies of the behavior of meteorites had led the astrophysicists Lindemann and Dobson to conclude in 1923 that some regions of the earth's upper atmosphere must be considerably hotter than the stratosphere, at $-50°$C. Whipple immediately recognized here a possible, simple explanation for abnormal audibility—for if the temperature of some layer of the upper atmosphere exceeded ground level temperature, sound rays would be refracted back to earth in a ring surrounding the source. *Why* any part of the upper atmosphere should be hotter than the stratosphere was left unanswered in 1923.

Ozone absorption of solar ultraviolet radiation is now believed to cause high temperatures in

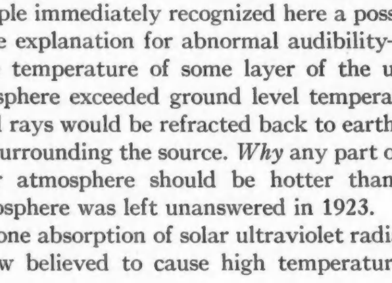

FIG. 1. Zones of audibility associated with an explosion in Kummersdorf, Dec. 18, 1925 (from Hergesell and Duckert).

[6] F. J. W. Whipple, *Nature* **111**, 187 (1923).
[7] V. I. Arabadzhi, *Meteorologia i Gidrologia* **5**, 21 (1946).

the air layer 30 to 60 km above the earth. Ozone is a most potent absorber of radiations having wavelengths between 2100A and 2900A. The highest concentration of ozone occurs about 21 km above the earth;[8] but, as pointed out by Penndorf,[9] maximum absorption could and probably does take place far above the level of maximum *possible* absorption. Furthermore, the content of ozone in the entire atmosphere is so small—condensed to standard conditions it would amount to a layer 3 mm thick—that in no sense need account be taken of the effect of this gas on the speed of sound except through its heating effect on the air with which it is mixed.[10]

Although most scientific evidence has added support to Whipple's explanation of abnormal sound, experiments performed northeast of Murmansk during the arctic night introduce perplexities. If the warm air layer in the upper atmosphere results from ozone absorption of solar radiation, no abnormal audibility zone should exist when the ozonosphere is continuously shadowed by the earth. Nevertheless, Wölcken's expedition recorded abnormal sounds during the winter solstice at 76° north latitude.[11] If abnormal audibility is caused by warm air in the ozonosphere, some new explanation must be found for the presence of warm air at high altitudes in the arctic shadow. Such would exist if meridianal transfer of air masses in the upper atmosphere were opposite in direction to that at low altitudes.[12]

Peacetime scientific use of V-2 rockets has recently supplied excellent confirmation of the heretofore more indirectly measured temperatures in the upper atmosphere. In the writer's opinion, the agreement removes all doubt from the validity of Whipple's theory, and firmly establishes abnormal sound measurements as a most useful method of upper atmosphere investigation.

[8] See F. W. P. Götz, *Ergeb. kosm. Phys.* **3**, 253 (1938) with 318 references; E. O. Hulbert, *J. Opt. Soc. Am.* **37**, 405 (1947).
[9] R. Penndorf, *Meteorol. Zeits.* **58**, 1 (1941), translated by C. C. Chapman, *Bull. Am. Met. Soc.* **27**, 331 (1946).
[10] Arabadzhi, reference 7, states that sound speed through the ozone layer would decrease because ozone is triatomic.
[11] K. Wölcken, *Zeits. f. Geophysik* **10**, 222 (1934).
[12] F. J. W. Whipple, *Quart. J. R. Met. Soc.* **61**, 285 (1935); W. J. Humphreys, *Physics of the air* (McGraw-Hill, 1940), p. 148.

Measurement Procedure

Upper atmosphere temperature values from abnormal sounds can be considered reasonably accurate only when well coordinated, recorded data are obtained in appreciable quantity. Although interrogations of residents near sites of accidental explosions have supplied the qualitative information on which the theories were formed, temperature calculations require accurate traveltime data, which do not come from questionnaires.

Planned destruction of ammunition dumps in Europe after World War I produced sound waves of such magnitude that they could be recorded and accurately timed by distant observers,[13] and thus be used for temperature calculations.[14] The reckoning process begins with determinations of incident angles of the abnormal sound rays as they strike earth.

Consider two observation stations at the same elevation separated by a radial distance dr, small compared with their mean distance r from a sound source. With ground temperature T_0 common at the two stations, the Laplace equation gives c_0 for ground-level sound speed. A sound signal from the source reaches the more distant station dt seconds later than it reaches the nearer station. If the sound wave traveled horizontally, we would find $\mathrm{d}r = c_0\mathrm{d}t$. If the sound ray is not horizontal, but strikes the stations at an incident angle i_0 (measured from the vertical), then $c_0\mathrm{d}t$ will be less than dr, and the angle of incidence can be found from the relation:

$$\sin i_0 = c_0\mathrm{d}t/\mathrm{d}r = c_0/(\mathrm{d}r/\mathrm{d}t). \qquad (2)$$

The expression dr/dt is called the *interval speed* of sound between the two neighboring stations; and for all angles of incidence, except 90°, the interval speed will be greater than c_0. Angles of incidence of sound rays in zones of abnormal audibility are generally greater than 70°.

Now if we neglect winds, and assume that air temperature changes only with altitude, sound ray paths are symmetrical from source to observer: that is, the ray which strikes earth at an incident angle i_0 is the ray emanating from the source at this same angle (absolute value). Anywhere along the ray path where air temperature T is known, we can compute the variable angle of incidence of the ray i by employing Snell's law and the Laplace equation (1):

$$\sin i/\sin i_0 = c/c_0 = (T/T_0)^{\frac{1}{2}}. \qquad (3)$$

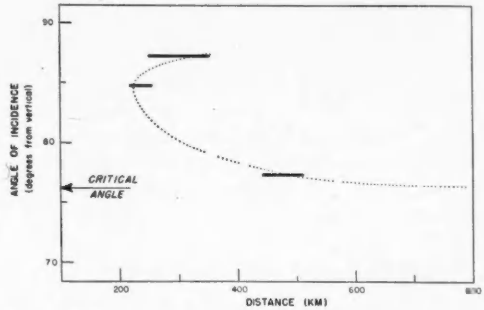

FIG. 2. Angles of incidence of sound rays refracted to earth from the ozonosphere. Solid horizontal lines are data "points" computed from interval velocities measured for the Helgoland explosion.

Comparing Eqs. (2) and (3), we obtain another useful relation. At the apex of the sound ray path, halfway between source and observer, the ray is horizontal ($i = 90°$). Therefore the interval speed determined between two adjacent stations at ground level equals the speed of sound at the apex of the ray striking between them. Air temperatures at the apexes of abnormal sound rays are thus obtained from measured interval speeds by means of the Laplace equation.

Learning the *altitudes* of abnormal sound-ray apexes is not so direct. The problem is very similar to that encountered in seismology, where, from measured traveltimes of longitudinal and shear waves, man has attempted to analyze the structure of the earth's crust. To the advantage of the air-seismologist comes one important factor: modern meteorological balloons can obtain soundings to altitudes far greater than the depths of the deepest wells. Assumptions need be made only for those altitudes *above* the summits of balloon flights.

High-altitude balloon flights made simultaneously with abnormal sound measurements supply the basic information for computing air temperatures at still greater heights. Starting

[13] G. Angenheister, *Zeits. f. Geophysik* **1**, 314 (1925).
[14] E. Wiechert, *Nach. Wiss. Göttingen* **1**, 49 (1925); B. Gutenberg, *Gerlands Beiträge z. Geophysik* **27**, 217 (1930); **35**, 46 (1932); F. J. W. Whipple, *ibid.* **31**, 158 (1931).

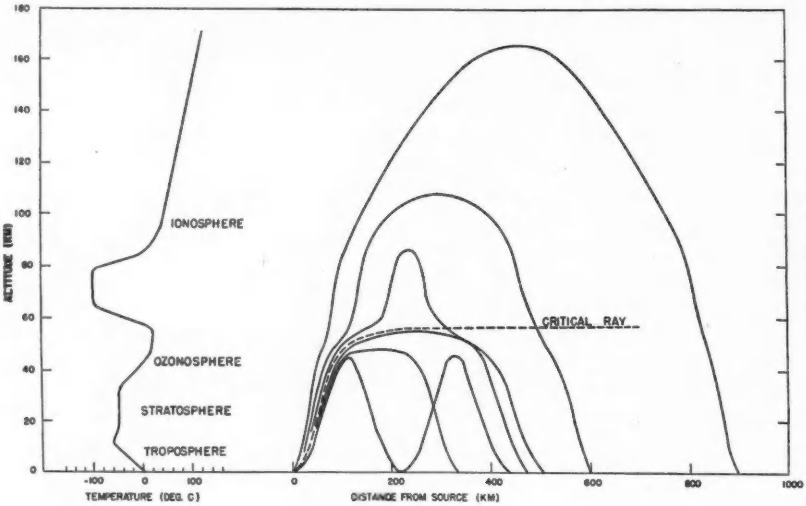

FIG. 3. Abnormal sound ray paths, and their associated temperature *vs.* altitude graph. Note the critical ray separating rays refracted to earth by the ozonosphere from those penetrating the ionosphere.

with measured angles of incidence at ground level, the ray paths and traveltimes are computed from ground to the top altitude reached by the balloons,[15] and ray paths are assumed symmetrical. By subtracting these *computed* traveltimes and travel-distances from *observed* traveltimes at a series of recording stations, one learns what horizontal distances the sound waves traveled, and what times were spent by the waves at elevations above the balloon summits.

Summarizing, we now know (*a*) the air temperature at the apex of each abnormal sound ray, (*b*) the incident angle of each ray as it enters the region above the balloon summits, (*c*) the horizontal distance between the points where each ray emerged from and re-entered the balloon-measured zone, and (*d*) the traveltime required for the sound wave to complete its arch. From this information we desire to learn just how high each sound ray arch should be constructed, because this knowledge combined with (*a*) will give us the temperature *vs.* altitude function in the upper atmosphere.

The problem of combining data (*b*), (*c*), and (*d*) to give heights of ray apexes was solved long ago by three mathematical physicists, Herglotz, Wiechert, and Bateman, primarily as a tool for seismologists. One assumption about the temperature-altitude function is required to make their variational equations applicable: the function must be monotonic in the region. When there is good reason to believe the function is not monotonic, the Herglotz-Wiechert-Bateman equations may be applied to sections of the problem.

An example of the calculation procedure is afforded by the latest published abnormal sound data, originating from the Helgoland blast,[16] at 1100 GCT, April 18, 1947. Angles of incidence of the abnormal sound rays, calculated from sound records obtained at a series of listening posts extending from northern Germany to northern Italy, are shown in Fig. 2. High-altitude meteorological balloons, released at blast time by a series of weather stations in Germany, succeeded in collecting temperature and wind velocity information to 29.5-km altitude. By applying the calculation processes outlined in the previous paragraphs, we arrive at the plot of abnormal sound ray paths, and the temperature *vs.* altitude function shown in Fig. 3.

Frequencies of Abnormal Sounds

The reader may well ask: "What kind of sound waves are these which travel at such

[15] A simple method using incomplete beta-functions has been derived by the author, *J. Acous. Soc. Am.* **19**, 832 (1947).

[16] S. W. Visser and J. Veldkamp, *Hemel en Dampkring* **45**, 150 (1947); E. F. Cox, J. V. Atanasoff, and B. L. Snavely, *Bull. Am. Met. Soc.* **29**, 78 (1948); E. F. Cox, *J. Acous. Soc. Am.* **20**, 549 (1948).

altitudes?" All abnormal rays shown in Fig. 3 have reached at least 42-km altitude en route. The simple barometric formula for ambient pressure P as a function of altitude h (km) is

$$P = P_0 \exp(-h/8). \qquad (4)$$

At 40-km elevation, ambient pressure is therefore about 5 mm of mercury or 6.6×10^{-3} atmosphere, and at 80 km ambient pressure is only 4.5×10^{-5} atmosphere. Sound cannot travel through a vacuum!

An eminent physicist, far better known for his equations in wave mechanics, treated this problem in his first major paper.[17] Absorption of sound waves in air depends on the ratio of sound wavelength to mean free path length of the gas molecules. Long wavelength, infrasonic pressure perturbations are easily transmitted through low-pressure gas wherein all audible frequency waves are absorbed. Indeed, Schrödinger's conclusions shown in Fig. 4 may well explain the persisting audible thump of the electric bell clapper in the bell jar demonstration (see reference 4, p. 155) long after reduced pressure has made inaudible the tinkle of the bell.

Abnormal sound signals are thus primarily inaudible, infrasonic waves, although low-bass audio rumbles may be heard. Most of the energy

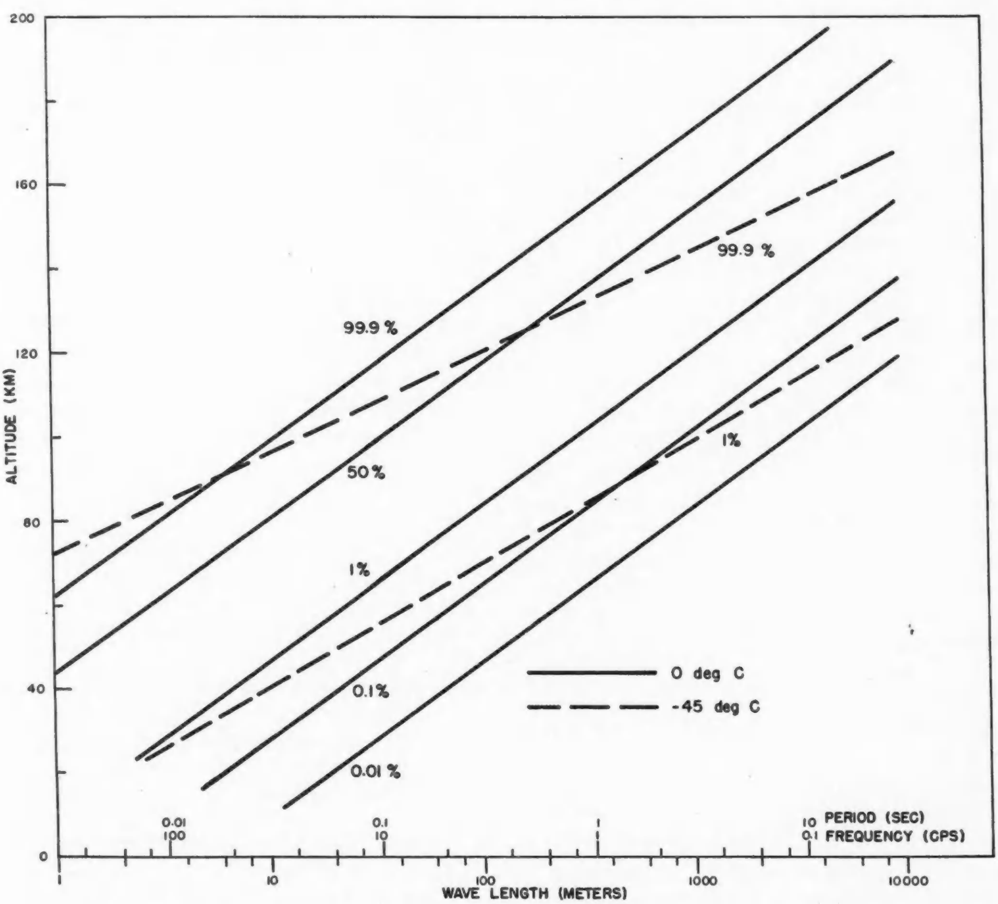

FIG. 4. Sound energy absorption per km of path in an isothermal atmosphere (after Schrödinger).

[17] E. Schrödinger, *Physikalische Zeits.* **18**, 445 (1917). See also R. B. Lindsay, *Am. J. Physics* **16**, 371 (1948).

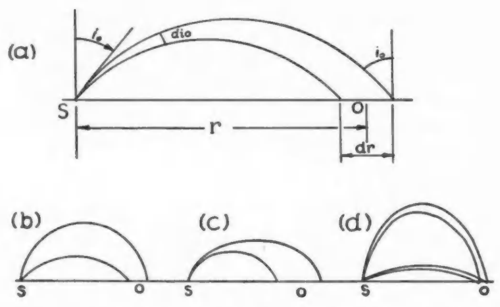

FIG. 5. Effect of the di_0/dr term on sound intensity. S is the source, O the observer (from Gutenberg).

is found in the frequency band $0.1-20$ cycle $\sec^{-1}$.

Condon has answered[18] the next question which the reader may propose. Is it proper to apply the Laplace equation (1), based on adiabatic compression and expansion, to pressure waves having periods as long as 10 seconds? The answer is affirmative. In explaining the error made by Newton, corrected by Laplace, many teachers make an error worse than that made by Newton. Isothermal expansion and adiabatic expansion are *not* to be differentiated by the speed at which expansion takes place. Sound waves are adiabatic waves because thermal gradients between pressure peaks and troughs are too small to grant isothermal conditions. Infrasonic waves of small amplitude are, therefore, truly adiabatic waves.

Infrasonic waves of the type treated in this paper by geometrical methods are not the lowest frequency, farthest traveling waves that have been observed. Super-colossal explosions, Krakatoa in 1883 and the Great Siberian Meteor in 1908, start large volumes of the atmosphere pulsating with periods of the order of 3 to 60 min. These oscillating pressure waves have been observed to encircle the earth[19] at a speed of about 310 m/sec. A wave-group speed of this low value can be derived from considerations of normal modes of oscillation of the atmosphere.[20]

[18] E. U. Condon, *Am. Physics Teacher* **1**, 18 (1933).
[19] A still earlier, small scale explosion, April 19, 1775, has so far been neglected by scientific writers. See R. W. Emerson "Concord Hymn," Louis Untermeyer, *A treasury of great poems* (Simon and Schuster, 1942), p. 791.
[20] G. I. Taylor, *Proc. Roy. Soc.* **A156**, 318 (1936); M. V. Wilkes and K. Weekes, *ibid.* **192**, 80 (1947); C. L. Pekeris, *Physical Rev.* **73**, 145 (1948).

Amplitude of Waves

Very little apparent relation exists between the sound amplitude observed in the abnormal zone and the magnitude of the explosion initiating the waves.[15] A real relation undoubtedly exists, but as the following derivation[21] clearly shows, meteorological factors far outweigh all others.

Let us assume an even distribution of sound energy over the upper hemisphere covering a blast at radial distance a ($a \approx 1$ km). Let W be the total sound energy in the hemisphere. The energy passing through an elemental ring of width di_0 on the sphere, at colatitude angle i_0, will be

$$dW = ((2\pi a \sin i_0 \, a di_0)/(2\pi a^2))W$$
$$= W \sin i_0 di_0. \quad (5)$$

After penetrating the upper atmosphere, this energy may strike an earth area $2\pi r dr$ at a great distance r from the blast (Fig. 5). Thus in the zone of abnormal audibility, the surface density of energy, neglecting absorption along the route, will be

$$\frac{W \sin i_0 di_0}{2\pi r dr} = \frac{W}{2\pi} \frac{\sin i_0}{r} \frac{di_0}{dr}. \quad (6)$$

The distance r at which abnormal sound rays strike the earth and the angle i_0 at which they strike are fixed by meteorological conditions. Near the inner boundary of the abnormal audibility ring, the sound waves leaving the source nearly horizontally come to earth with considerable strength.[22] The graph of i_0 *vs.* r, Fig. 2, predicts very high intensity near 220 km.

Effects of the di_0/dr term of Eq. (6) are illustrated in the lower parts (b, c, d) of Fig. 5. When the energy contained in a relatively large incremental angle di_0 strikes a narrow ring, as in (b), high intensity results. On the other hand, if the angle of incidence is changing very slowly in the neighborhood of a receiving station, as in case (c), low intensity is observed. Case (d) shows how, under exceptional meteorological conditions, a *normal* sound signal held near the

[21] Essentially that given by R. Emden, *Meteorol. Zeits.* **35**, 114 (1918), and by B. Gutenberg, *J. Acous. Soc. Am.* **14**, 151 (1942).
[22] Abnormal sound waves from Pacific Fleet gunfire off Santa Catalina Island have shattered windows in Bakersfield, although no sound was heard in Los Angeles (B. Gutenberg and C. F. Richter, *Gerlandik Beitr. z. Geophysik* **31**, 155 (1931).

earth by an inversion layer in the troposphere, and an *abnormal* signal traveling the upper path could produce equal perturbation pressures at a distant recording station. A microbarographic record of this unusual phenomenon is shown as Fig. 6.

Separation and multiplicity of abnormal audibility zones, illustrated by Fig. 1, are also closely related to the term di_0/dr of Eq. (6). The ozonosphere has at any given time a very definite maximum temperature T_m. This maximum temperature establishes a critical angle of incidence $(i_0)_c$ at the source, whose value can be calculated by the relation

$$\sin(i_0)_c = (T_0/T_m)^{\frac{1}{2}}. \tag{7}$$

A sound ray leaving the source at the critical angle of incidence is refracted to infinity at the altitude of the hottest layer of the ozonosphere (see Fig. 3). A ray whose starting angle is slightly larger than critical returns from the ozonosphere at a finite distance from the source. Thus a trivial amount of sound energy is spread over a very large area, di_0/dr is essentially zero as the critical angle is approached, and the first zone of abnormal audibility has a practical outer boundary. Figure 2 shows 500 km as the outer boundary for instrument recording at the time of the Helgoland blast. The first abnormal zone is not nearly so wide for aural observations.

Second and third zones of abnormal sound result from reflection of the waves by the earth. Traveltimes to outer zones are multiples of the traveltime from source to the first abnormal zone.

Whipple's explanation of the abnormal sound phenomenon has several times[2, 7] been questioned on the basis of surmised abnormal sound pressures. If in the stratosphere or upper atmosphere, the sound wave perturbation pressure p is an appreciable fraction of the ambient pressure P, the Laplace equation (1), valid for an infinitesimal ratio of p/P, cannot be expected to hold. In sound waves of finite amplitude, compressions travel with supersonic speed and rarefactions proceed with subsonic speed. An N-shaped wave results, whose steep front travels at a supersonic speed V given by the equation

$$V = c[1+\beta(\gamma+1)/2\gamma]^{\frac{1}{2}}, \tag{8}$$

where β is the pressure ratio p/P.[23] If β increases with altitude along a wave front, the wave will be refracted downward. Might this shock-wave refraction process not explain abnormal audibility zones?

To test this proposition, let us first examine the approximate shape of the wave front after it has all passed through the tropopause of a typical atmosphere. Let ground temperature be 10°C, and the isothermal layer temperature be −60°C everywhere above 12 km. On emerging from the tropopause at a horizontal distance 46 km from the source, the ray which started horizontally from the source ($i_0 = 90°$) has an incident angle 60°. The ray which started at an angle of inci-

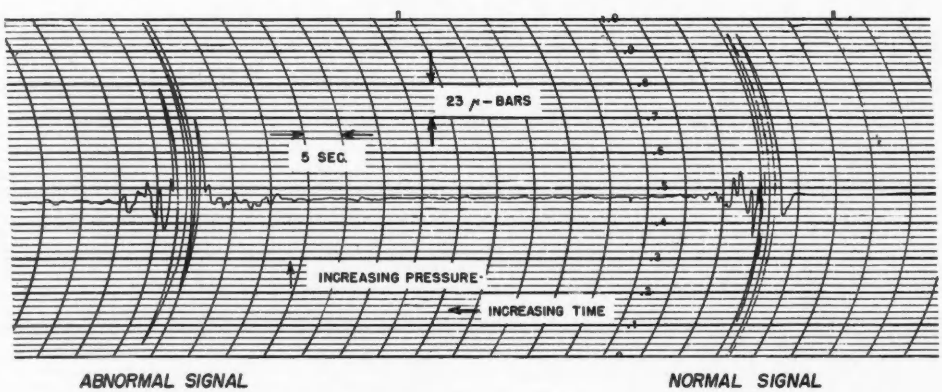

FIG. 6. Microbarograph record of sounds received 182 km away from a single 125-ton TNT blast.

[23] A simple derivation is given by J. W. M. DuMond *et al.*, *J. Acous. Soc. Am.* 18, 97 (1946). Their Eq. (10), p. 104, contains a misprint ρ_0 for p_0.

dence of 70° enters the stratosphere at a point half as far away (horizontal distance), and its inclination is 54.5°. Since the initial vertical divergence has been reduced from 20° to 5.5°, it is highly improper to regard the lower section of the wave front in the stratosphere as being spherical about the source. The divergence between the rays starting from the source at 90° and 80°, respectively, is reduced to only 1.5° in the stratosphere. The lower section of the wave front is therefore essentially conical, with its normal 60° from the vertical.

Along the vertical cross section AB of the lower part of the wave front (Fig. 7) the acoustic intensity, $p^2/\rho c$, is constant. In the isothermal stratosphere, c has the same value at points A and B, but ρ decreases exponentially according to the equation

$$\rho_B=\rho_A \exp(-L \sin i/8), \tag{9}$$

where L is the length of the wave-front segment AB measured in kilometers. We thus find perturbation pressures at points A and B related by

$$p_B=p_A \exp(-L \sin i/16), \tag{10}$$

whereas ambient pressures at the two points, from Eq. (4), have a ratio $\exp(-L \sin i/8)$. The ratio of perturbation pressure to ambient pressure at point B is greater than that at point A according to the equation,

$$\beta_B=\beta_A \exp(L \sin i/16), \tag{11}$$

and, as a consequence, point B has a greater forward speed than point A.

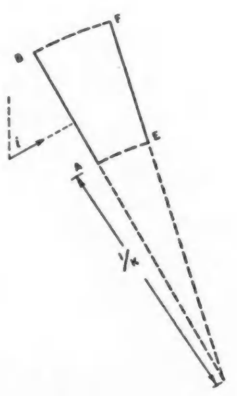

FIG. 7. Refraction of a shock-wave in the isothermal stratosphere.

Curvature of the wave normal is now easily found, because the wave front moves from B to F while another part of it moves from A to E. If K be the curvature,

$$K=((V_B/V_A)-1)/(L). \tag{12}$$

Expanding Eq. (8) and rejecting terms higher than first order in β, shows

$$V=c[1+\beta(\gamma+1)/4\gamma], \tag{13}$$

which for air, wherein $\gamma=1.40$, becomes

$$V=c[1+0.43\beta]. \tag{14}$$

By substitution in Eq. (12) we find

$$K=0.43(\beta_B-\beta_A)/L=0.43\beta_A \sin i/16(\text{km}^{-1}). \tag{15}$$

For sound rays to return to earth, i must change from 60° to 90° in the stratosphere. The radius of curvature therefore ranges between $37/\beta$ km and $43/\beta$ km.

Now Wegener[2] has found that the inner boundary radius of the first zone of abnormal audibility depends on the season, and ranges between 105 km in winter, 200 km in summer. For sound waves to return to earth at such distances as a result of the shock-wave process, β must be of the order of 0.3 or greater. If the essentially cylindrical divergence computed for the inner 50 km of horizontal travel prevailed for the remainder of the path, we might expect to observe peak perturbation pressures of the order of 4×10^3 dyne cm^{-2} in the first zone of abnormal audibility. Measurements reported by Cox[15] show actual pressures to be less than 3 percent of this value.

Furthermore, Eq. (15) shows that if the shock-wave process were the cause of abnormal audibility, the distance from an explosion to the abnormal sound ring should decrease as the size of explosion increases. Explosions varying in size all the way from an 8-inch gun blast[12] to an atomic bomb[24] seem to produce abnormal sound zones at essentially equivalent distances from the sources. Although the inner radius of the first zone of abnormal audibility does vary with the season,[2] and may depend on latitude, it appears to be *completely independent* of blast size.

With the shock-wave phenomenon thus elimi-

[24] B. Gutenberg, *Bull. Seismolog. Soc. Am.* **36**, 327 (1946).

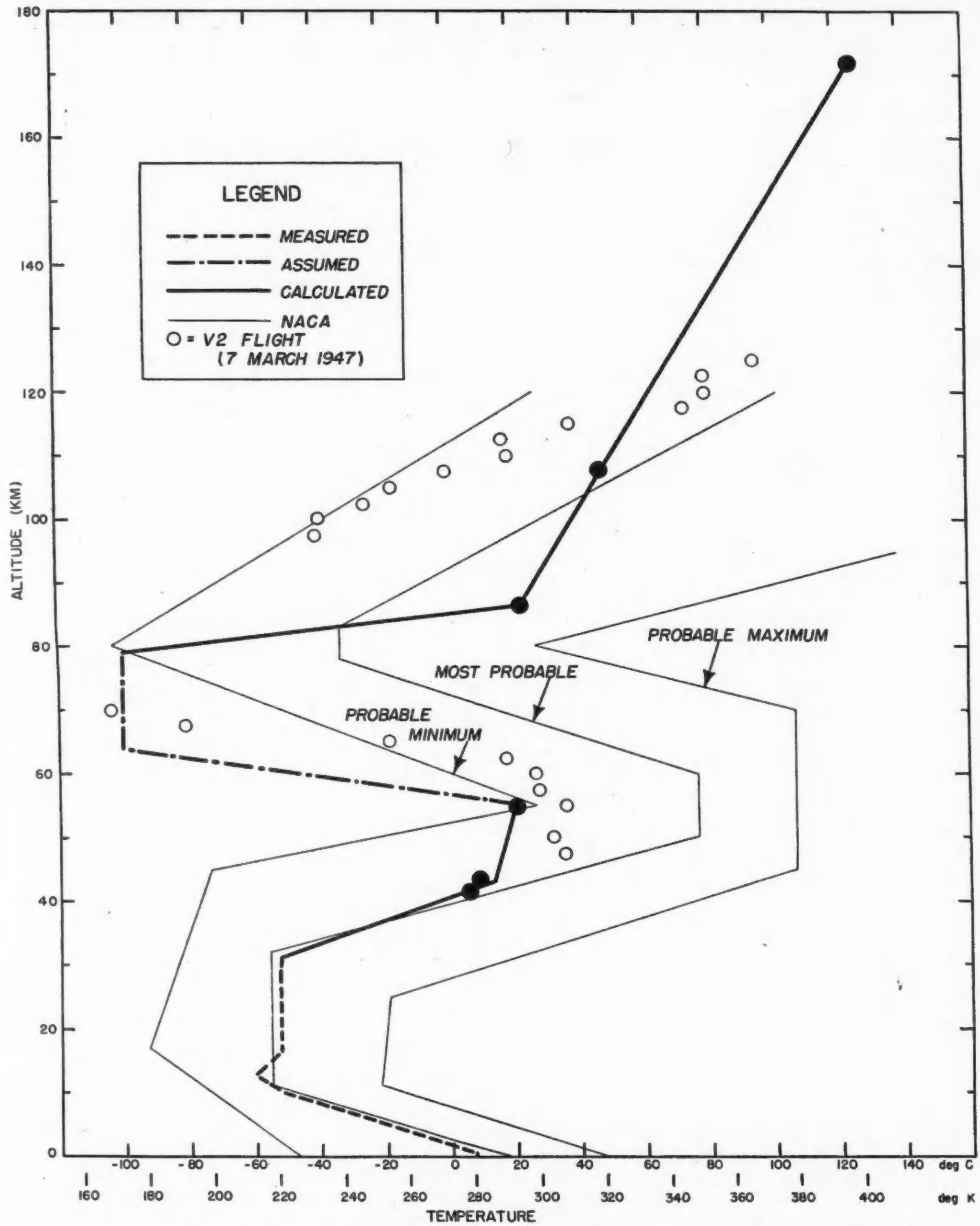

FIG. 8. Comparison of measured upper atmosphere temperatures (Helgoland blast), V-2 rocket results, and NACA tentative standards.

nated as a possible explication of abnormal sound on the bases of pressure measurements and lack of variation with size of explosion, Whipple's theory remains the most logical explanation.

Conclusions

To date no truly *direct* method of measuring temperatures in the upper atmosphere has been devised. Perhaps the measurements employing V-2 rockets[25] are less indirect than abnormal sound measurements, but the high cost of V-2 rocket experimentation severely limits the quantity of results to be expected. The cost of gathering abnormal sound data satisfactory for calculation of meteorological factors is not trivial, but many investigations could be made for the price of one V-2 rocket and launching.

The National Advisory Committee for Aeronautics Special Subcommittee on the Upper Atmosphere has considered all types of measurements made prior to 1947, and constructed therefrom a "tentative standard atmosphere" to 120-km altitude.[26] "Most probable," "probable minimum," and "probable maximum" values have been given.

Temperatures calculated from abnormal sound measurements on the Helgoland blast, and from V-2 data,[27] are compared with the NACA tentative standards in Fig. 8. The writer has very little confidence in the abnormal sound results above 110-km altitude, and also questions the superadiabatic lapse rate between 55 and 65 km.

Unfortunately, reduction of both V-2 data and abnormal sound data to temperature values makes use of the Laplace equation (1). There are good reasons to believe that above about 80 km a considerable portion of the oxygen is atomic rather than molecular. Should this be true, the molecular weight of the air above 80 km is between 29 and 24, and γ is between 1.40 and 1.46. Temperatures calculated on the basis of ground level air composition, such as those of Fig. 8, excepting the NACA values, may therefore be as much as 26 percent too high (Kelvin scale). Correction of temperatures must await more knowledge of upper atmosphere composition.

In the meantime, much of interest remains to be learned of the lower altitudes, particularly the ozonosphere. No experimental study of diurnal temperature variation has been published; and the facts concerning ozonosphere temperature dependence on latitude are too scant to substantiate existing theories. Carefully organized experiments with abnormal sound could supply this information expeditiously.

Acknowledgments

The author wishes to express his gratitude to his colleagues in the Acoustics Division of the Naval Ordnance Laboratory who were responsible for the design and construction of the apparatus and equipment used in obtaining the data presented in this paper. They also assisted materially in the acquisition and reduction of those data.**

[25] N. Best, R. Havens, and H. LaGow, *Physical Rev.* **71**, 915 (1947).

[26] NACA Tech. Note No. 1200 (January 1947); *Bull. Am. Met. Soc.* **28**, 425 (1947).

[27] Recalculated data presented by R. J. Havens at the Chicago meeting of the Am. Physical Soc. (December, 1947).

** The basic material presented in this paper will appear shortly in the *Journal of Meteorology* in an article entitled "*Upper atmosphere temperatures from Helgoland big bang*," by E. F. Cox, J. V. Atanasoff, B. L. Snavely, D. W. Beecher, and J. Brown, all staff members of the Naval Ordnance Laboratory.

Errata

In the note, **Meaning of the Ratio e/m**, by Robert Weale, *Am. J. Physics* **16**, 358 (1948), and in the note, **The Ratio e/m**, by P. Le Corbeiller, *Am. J. Physics* **16**, 358 (1948) the name Robert Weale was incorrectly printed as Robert Weil.

In the article, **Concerning Rope Problems and the Principles of Momentum and Work,** by Gordon Ferrie Hull, *Am. J. Physics* **16**, 447 (1948), lines 10 and 11, column 2, p. 448, should read

$$\frac{\mathrm{d}p}{\mathrm{d}t}=\mathrm{d}[m_0v/(1-\beta^2)^{\frac{1}{2}}]/\mathrm{d}t=\mathrm{d}[m_0c\beta/(1-\beta^2)^{\frac{1}{2}}]/\mathrm{d}t.$$

Radiation from Rocket Flames and Its Effect on Rocket Performance*

S. S. Penner
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

THE development of modern jet propulsion devices[1] entails the solution of numerous special problems, one of which, the effect of radiation on rocket performance, forms the subject of our discussion. By rocket performance is meant combustion performance rather than behavior as a projectile. In order to evaluate the effect of radiation on rocket performance it is, of course, necessary to determine first the nature of the radiation emitted from rocket flames—that is, to measure the radiant intensity and its dependence on wavelength as well as its variation with temperature. After the nature of the radiation from rocket flames has been ascertained, it is possible to consider in detail a number of problems that may be encountered in combustion processes occurring in chambers other than a rocket motor. An exhaustive discussion of all of the manifold effects of radiation on rocket performance falls outside the scope of this article. Two important problems which are not considered in detail are the following: (a) the chemistry of combustion and its dependence on the presence of activated molecules which may be formed by photochemical processes; (b) radiant heat transfer to a rocket chamber, although it is known that radiant heat transfer may amount to 50 percent or more of the total heat transfer in rocket motors utilizing hot propellant systems such as hydrogen and oxygen or hydrogen and fluorine. In view of the difficulty of obtaining refractory materials or metals capable of withstanding the high temperatures (sometimes exceeding 3000°K) encountered in rocket chambers, the practical importance of radiant heat transfer to the motor wall is obvious.

The present review is primarily concerned with (a) a discussion of the experimental evidence relating to the nature of the radiation from rocket flames and (b) a survey of the theoretical and experimental results relating to the manner in which radiant energy is absorbed by solid propellant grains whose rate of burning and behavior are changed because of the absorption of radiant energy. Solid-fuel propellant grains, which burn in a direction normal to their exposed surface, are encountered in one large class of rockets of which the jet-assisted-take-off units are representative examples.[1]

In connection with the discussion of the nature of the radiation from rocket flames, a brief description is given of the two-color pyrometric technique for determining flame temperatures. The meaning of temperature measurements in nonequilibrium systems is considered. As regards the effect of radiation on the burning processes inside a rocket chamber (interior ballistics), it is shown how the transfer of radiant energy to a burning propellant grain may be responsible for erratic burning associated with radiation fissuring caused by excessive heating in the interior of a powder grain. In agreement with expectations, radiation fissuring can be largely eliminated by decreasing the amount of radiant energy which can be transmitted through a burning propellant grain to its interior. The relation between rate of burning of a solid propellant grain and the absorption of radiant energy is considered. Finally, the effect of radiation on performance is evaluated with respect to work currently in progress on combustion problems in other fields.

Nature of the Radiation from Rocket Flames

The combustion processes in liquid-fuel and solid-fuel rockets take place with the evolution of radiant energy. The energy distribution as a function of wavelength is dependent upon the chemical species present during combustion. Since the chemical reactions in a rocket chamber may involve the temporary formation of energy-rich free radicals and other transient species,[2] the

* Presented before the *Symposium on heat-transfer problems in jet propulsion devices*, Pasadena, California, June 24, 1948.

[1] The physics of rockets has recently been discussed in a series of three papers by Seifert, Mills, and Summerfield (*Am. J. Physics* **15**, 1–21, 121–140, 255–275 (1947)). Reference should be made to these articles for a description of the general characteristics of rockets.

[2] For a review of combustion theories see, for example, W. Jost, *Explosion and combustion processes in gases* (translated by H. O. Croft, McGraw-Hill, 1946); B. Lewis and G. von Elbe, *Combustion flames and explosions of gases* (Cambridge Univ. Press, 1938).

radiation emitted from reacting propellant gases would be expected to be a complicated function of time and position.[3–5] It follows, therefore, that accurate calculation of radiant heat transfer to rocket tubes, to burning solid propellants, and to moving liquid droplets entails all of the difficulties of the solution of a complex problem in chemical kinetics.

The determination of the nature of the emitted radiation is further complicated by the lack of data relating to the emission characteristics of mixtures of gases which may contain solid particles or liquid droplets. This lack of data is not surprising since even the emission characteristics of homogeneous gaseous systems in complete thermal equilibrium can be only approximated. Problems of (a) pressure broadening of spectral lines and of (b) perturbations caused by the presence of foreign gases and by the electric fields of ions and free radicals inevitably enter into the calculation of radiant heat transfer in rockets.[6] In fact, the entire problem is so hopelessly complicated that drastic simplifications are required before any estimates of orders of magnitude of radiant heat transfer can be made.

Since a wide variety of chemical species co-exists at elevated temperatures and pressures with solid or liquid particles in many rocket motors, it is not unreasonable to expect a more or less continuous distribution of radiant energy as a function of wavelength. Considerable experimental evidence exists[3,4] for the presence of a continuous background of radiant energy. Furthermore, it appears[3,4] that the radiant energy distribution in some solid-fuel rocket motors operating at pressures of about 2000 lb/in^2 is similar to the radiant energy distribution produced by blackbody or graybody emitters. There is, however, no evidence for graybody emission characteristics of the gases in liquid-fuel rockets, which are usually operated at a pressure of only a few hundred pounds per square inch. For the larger liquid-fuel service rockets an increase in radiation path length would tend to compensate for a decrease in the chamber pressure.

Since the assumption of graybody characteristics for the emitters in rocket chambers is made in all of the calculations relating to the effect of radiation on rocket performance, the experimental evidence in support of this assumption will be reviewed briefly in the succeeding section. Calculations of the effect of radiation on rocket performance will be described later in the article.

Experimental Evidence Relating to the Nature of the Radiant Energy Emitted from Flames of Double-Base Powders in Solid-Fuel Rockets

Some exploratory work on the nature of the radiant energy emitted by the propellant gases issuing from double-base powder charges was published in France and England during the early years of World War II. A double-base propellant is a solid mixture of combustible material consisting of approximately 40 percent of nitroglycerin and 60 percent of nitrocellulose, in which some of the nitroglycerin or nitrocellulose may be replaced by other substances in order to improve the burning characteristics of the powder. Muraour and Aunis[7] concluded from a comparison between the spectra of powder gases and the continuous radiation from a carbon arc that the gases radiated as a blackbody. Similarly, Price and Norrish[8] and Price and Philpotts[9] observed a continuous spectrum together with lines of Ca, K, and Na, and bands of CaO during the burning of cordite in closed vessels.

In this country the most extensive studies relating to the nature of the radiation from powder gases were made by Kracek and Benedict[3] at the Geophysical Laboratory, Washington, D. C. These investigators studied the radiation from powder gases in closed chambers by photographic and spectroscopic methods. They also developed

[3] F. C. Kracek and W. S. Benedict, *An experimental study of powder gas radiation and temperature*, OSRD Report No. 3291, February 1944.

[4] R. S. Craig, OSRD Report No. 5832, December 1945.

[5] D. R. Bellman (personal communications) has clearly demonstrated the nonstationary character of the combustion pattern by studies with transparent rocket motors.

[6] For a review of radiation from perturbed systems, see H. Margenau and W. W. Watson, *Rev. Mod. Physics* **8**, 22 (1936). Some more recent articles on the pressure broadening of spectral lines are the following: A. Jablonski, *Physical Rev.* **68**, 78 (1945); H. M. Foley and D. M. Dennison, *Physical Rev.* **61**, 386 (1942); H. M. Foley, *Physical Rev.* **69**, 616 (1946).

[7] H. Muraour and G. Aunis, *Chaleur et Industrie* **20**, 31 (1939).

[8] W. C. Price and R. W. G. Norrish, Adv. Council Sci. Res. Tech. Dev. to British Ministry of Supply, A. C. 3410.

[9] W. C. Price and A. R. Philpotts, Adv. Council Sci. Res. Tech. Dev. to British Ministry of Supply, A. C. 3870; A. C. 4031.

a two-color photoelectric method for the measurement of flame temperature, mass emissivity, and absorptivity of powder gases in closed chambers and in guns. The two-color photoelectric pyrometer was later applied by Craig[4] to the measurement of temperatures and to the approximate determination of emissivities of powder gases in solid-fuel, double-base propellant rockets operating under conditions similar to those encountered in service units.

Kracek and Benedict[3] found that when the gas densities in their closed chamber exceeded about 0.03 g/cm^3 for a radiation path length of 2 cm, then the observed spectra showed only continuous radiation with blackbody characteristics. At lower densities they observed continuous radiation with line and band structure corresponding to the inorganic constituents of the powders. These observations do not appear to be in disagreement with some results obtained by Daniels and co-workers[10] showing relatively intense vibrational bands of CO_2 and H_2O in the near-infra-red spectra of double-base powders, since Daniels' measurements were made at lower pressures with thin layers of emitter. Beek[11] appears to favor the view that propellants containing a few percent of inorganic matter (most double-base rocket propellants fall in this category) radiate with an energy dependence on wavelength corresponding to the radiation from a graybody. He assigns a mass emissivity of the order of 100 cm^2/g to conventional double-base rocket propellants where the mass emissivity k is defined by the relation[12]

$$I = I_0(1 - e^{-k\rho L}). \quad (1)$$

Here I, in ergs/(cm^2 sec), is the radiant energy over all wavelengths actually emitted from 1 cm^2 of surface of a gaseous layer L cm thick and of density ρ g/cm^3, if the radiant energy emitted over all wavelengths by a blackbody at the same temperature as the radiating gases is I_0 ergs/(cm^2 sec). The value of 100 cm^2/g for the mass emissivity was chosen on the basis of some results obtained by Kracek and Benedict[3] which had led these authors to conclude that the source of the continuous radiation is finely divided liquid or solid particles. The extensive measurements of the color temperature of powder gases in rockets operating at a pressure of a few thousand pounds per square inch made by Craig[4] do not provide independent confirmation of the hypothesis that the emitters are blackbodies or graybodies. In fact, Craig's emissivity measurements seem to indicate that the emissivities for the two wavelength regions selected by him may differ by a factor of 2 or 3. These results will be considered more fully later.

Since the most important experimental studies of radiant energy from rockets are based on application of two-color pyrometry, it seems desirable to consider the meaning of temperature in nonequilibrium systems and then to describe the two-color technique in some detail. First, however, some interesting observations incidental to the principal program of radiation studies should be pointed out because these results do not appear to have received the attention which they deserve in connection with current studies of the combustion mechanism.

Kracek and Benedict,[3] Craig,[4] and Bellman[5] have obtained experimental results which clearly demonstrate the existence of nonstationary flame patterns. Rapid fluctuations in temperature and emissivity appear to be a part of the "normal" combustion process. Kracek and Benedict[3] observed pressure and light intensity variations at a frequency of about 2000 c/sec in closed chambers and suggested that these oscillations behave as adiabatic compressions and rarefactions and are therefore similar to standing sound waves. Craig[4] observed color temperature variations of several hundred degrees and emissivity fluctuations of about 20 percent in the solid-fuel rocket motors studied by him. He suggested that these fluctuations were caused by the turbulent flow of hot and cool masses of gas before the observation window or by the shifting of position of standing shock waves.

Since a considerable amount of work is now being done on the mechanism of combustion, it appears likely that more detailed information regarding the nature of the radiant energy emission from flames will be forthcoming as the result of these studies. Examples of this type of work

[10] F. Daniels *et al.*, OSRD Report No. 3206, 1944.

[11] J. Beek, Jr., OSRD Report No. 4033, August 1944; Part I of OSRD Report No. 5817, 1946.

[12] For a discussion of radiant heat-transfer calculations and numerous literature references, see H. C. Hottel's chapter on radiant heat transfer in W. H. McAdam's, *Heat transmission* (McGraw-Hill, 1942).

are experiments on low-pressure flames now in progress at the Jet Propulsion Laboratory[13] and studies of emission spectra at the Applied Physics Laboratory.[14]

The Meaning of Temperature in Nonequilibrium Systems and the Two-Color Technique for Measuring Flame Temperature

A discussion of the meaning of temperature in nonequilibrium systems can easily lead to philosophic arguments which it appears desirable to avoid here.[15] From the operational point of view, there are as many different temperatures as there are methods for measuring temperature. Thus, we may speak of a rotational temperature determined from a variation in intensity of the emitted radiation corresponding to a change in rotational quantum number and obtain different numerical results for each molecular species selected for measurement; we may speak of a temperature defined through the temperature of line reversal of a given resonance line such as one of the lines of the well-known doublet of sodium; or we may speak of a thermocouple temperature measured by the use of a balancing circuit and a bimetallic element and obtain results strongly dependent on the choice of thermocouple material. It is apparent that none or all of these definitions of temperature may be significant, depending upon their agreement with one another and with a criterion which is known to be useful for indicating combustion efficiency.

In the case of combustion in a rocket chamber, we are primarily interested in obtaining the maximum possible amount of useful work from a given system of chemical reactants. A measure of the actual useful work obtained from a propellant system is the specific impulse, the definition and measurement of which have been described in the literature.[1] Closely related to the maximum possible work obtainable is the thermodynamic equilibrium temperature reached in the combustion chamber if complete chemical equilibrium is attained with respect to all possible reactions and no heat is lost from the reacting propellants. Experience has shown that the numerical values of temperature obtained by a given set of measurements approach but never exceed the thermodynamic equilibrium temperature even in nonequilibrium systems, provided the method whereby the temperature is determined involves measurements on quantities which are in statistical equilibrium with the translational energy distribution of the molecules. We may therefore regard a value of the temperature determined under conditions of statistical equilibrium as a direct measure of combustion efficiency. In practice it may prove to be exceedingly difficult to eliminate all instances of abnormal excitation which would lead to erroneous results. For this reason it is preferable to make temperature determinations by several independent methods rather than to assume that any one procedure will lead to a significant value for the flame temperature. If two or more independent experimental techniques lead to the same numerical result for the flame temperature in nonequilibrium systems, then it appears very likely that the true flame temperature (defined through the random translational speed of the gas molecules) has been measured. The two-color pyrometric technique described below would probably yield a satisfactory temperature measurement (assuming that the emissivity is known as a function of wavelength) even if some of the emitters in the wavelength region over which measurements are made were preferentially excited, as long as the fraction of such emitters remains negligibly small compared with the total number of emitters. On the other hand, it is evident that the two-color pyrometric technique for measuring flame temperatures gives average values over a relatively extended region of the flame zone and that therefore the relation between the measured temperature and the translational temperature at a given point is not obvious.

The rate of emission of radiant energy from a blackbody is given by the well-known relation derived by Planck in 1900. Planck's equation reduces to Wien's equation for all wavelengths less than about 1μ if the temperature is approximately 3000°K. Wien's equation, applicable as a very close approximation, is

$$J(\lambda, T)d\lambda = (C_1/\lambda^5)\exp(-C_2/\lambda T)d\lambda, \qquad (2)$$

[13] M. Gilbert, private communication.

[14] S. Silverman, G. A. Hornbeck, and R. C. Herman, *J. Chem. Physics* 16, 155 (1948).

[15] For a review of techniques used for flame temperature measurements, see B. Lewis and G. von Elbe, *Temperature, its measurement and control in science and industry* (Reinhold, 1941), pp. 707–719.

where $J(\lambda, T)d\lambda$ is the radiant intensity in ergs/(cm^3 sec) emitted by unit area of a blackbody radiator over a solid angle of 2π-steradians at the temperature T in the wavelength region between λ and $\lambda+d\lambda$, $C_1=3.74\times10^{-5}$ erg cm^2/sec, and $C_2=1.432$ cm °K.

The spectral emissivity $\epsilon(\lambda, T)$ is defined as the ratio of the radiation intensity emitted by a given substance to the radiation intensity emitted by a blackbody at the same temperature. The color temperature T_c of an emitter is defined as the temperature at which a blackbody emits radiation having the same ratio of radiant intensities at the wavelengths λ_1 and λ_2 as the emitter under study. This definition of T_c can be expressed analytically by the relation

$$\frac{J(\lambda_1, T_c)}{J(\lambda_2, T_c)}=\frac{\epsilon_1(\lambda_1, T)J(\lambda_1, T)}{\epsilon_2(\lambda_2, T)J(\lambda_2, T)}. \tag{3}$$

Combining Eqs. (2) and (3) leads to the following relation between color temperature T_c and true temperature T:

$$(1/T)-(1/T_c) = [\ln(\epsilon_1/\epsilon_2)]/C_2[(1/\lambda_1)-(1/\lambda_2)]. \tag{4}$$

It is evident from Eq. (4) that the color temperature will be equal to the true temperature, independent of the choice of the arbitrary wavelengths λ_1 and λ_2, if the substance under study is a blackbody ($\epsilon_1=\epsilon_2=1$) or a graybody ($\epsilon_1=\epsilon_2<1$). Equation (4) suggests a very direct test for the existence of blackbody or graybody emitters in the powder gases formed from a propellant of given composition. The right-hand side of Eq. (4) will not be constant, in general, for different values of λ_1 and λ_2. However, if the observed color temperatures are independent of the arbitrary wavelengths λ_1 and λ_2, then it is reasonable to conclude that the right-hand side of Eq. (4) vanishes and that therefore blackbody or graybody emitters are present. Kracek and Benedict[3] have used this criterion, among others, to demonstrate the existence of blackbody emission in their closed chamber studies by replacing the wavelengths λ_1 and λ_2 by wavelength regions of finite widths. A similar experiment has not been carried out for solid-fuel rocket motors.

In Craig's[4] studies of the radiation from rockets, he observed the radiation normal to the chamber axis through quartz windows of the type introduced by Poulter.[16] The radiation was interrupted 240 times per second before entering a photo-tube box in which the light was split into two approximately equal components by passage through a semitransparent mirror aligned at an angle of 45° with the axis of the beam. One component of the incident light fell on a red-sensitive photo-cell after passing through suitable filters. The red photo-cell-filter combination responded to radiation between 0.70μ and 1.2μ only. The other component of the incident light was similarly received by a blue-sensitive photo-cell-filter combination which responded to light of wavelengths between 0.31μ and 0.60μ only. The photocell output consisted of pulsating direct current, the oscilloscope trace of which was photographed after suitable amplification. By making proper allowance for the light response of the photo-cell-filter combination, it was then possible to determine the light intensity emitted by the rocket in the selected red and blue wavelength regions. The apparatus was calibrated by comparison with a tungsten strip lamp of known emission characteristics.

Craig's measured color temperatures are equal to the actual flame temperatures only if the emitters of radiation are graybodies. The presence of graybody emitters has not been proved independently by Craig. Instead, interpretation of his results is based on the work of Kracek and Benedict. Therefore, since the radiation path length of 6.5 cm in Craig's experiments was roughly 2.5 times that of Kracek and Benedict[3] in their closed chamber studies, there are no reasons for assuming that the emitters were blackbodies below a density of about 0.012 g/cm^3, corresponding to a pressure of about 1500 lb/in^2. Thus there are also no reasons for supposing that color temperatures and flame temperatures are equal at pressures below about 1500 lb/in^2. Craig also studied the rocket flames with a quartz spectrograph and obtained results similar to those observed by Kracek and Benedict, namely, a discrete line and band structure on a continuous background. The intensity of the continuous background increased with pressure, path length, and amount of solid material present. Finally,

[16] T. Poulter, *Physical Rev.* **40**, 860 (1932).

Craig made some direct measurements of emissivity in the blue and red spectral regions, using approximate calculated values for the flame temperature. A fourfold increase in emissivity was observed in some cases in going from an unsalted powder to a powder containing about 5 percent of K_2SO_4. Craig's emissivity studies indicate that the red and blue emissivities usually do not differ by a factor of more than 3.

The available experimental evidence for the existence of blackbody emission characteristics from the flames of double-base rocket propellants of the type used by Kracek and Benedict permits only the single quantitative assertion that the mass emissivity k must be sufficiently great to make the term $\exp(-k\rho L)$ negligibly small compared with unity if $\rho \geqslant 0.03$ g/cm^3 and $L \geqslant 2$ cm. The term $\exp(-k\rho L)$ is less than or equal to 0.01 for $\rho = 0.03$ g/cm^3, and $L = 2$ cm if k is greater than or equal to 77 cm^2/g. We may therefore regard the value $k = 77$ cm^2/g as an approximate lower limit for the mass emissivity of salted and darkened double-base rocket propellants. The assumed existence of graybody emission characteristics for all powders is probably also justified, as a first approximation. The approximation will be the better the larger the pressure, the longer the radiation path length, and the higher the concentration of solid or liquid particles in the flame. The fact that theoretical predictions based on the assumption of graybody emission characteristics are in good agreement with experimental results, as will be shown in the next section, may be viewed as *a posteriori* evidence for the validity of this assumption, as a first approximation, even when the effective mass emissivities are 20 to 40 cm^2/g as in some of the approximate direct measurements made by Craig.

The Effect of Radiation on Rocket Performance

Historically the development of theories for explaining the effect of radiation on rocket performance is connected with the selection of proper darkening agents for transparent double-base powders. Before the nature of the phenomenon was fully understood, various explanations, such as biochemical disintegration resulting in the production of wormholes, were offered for the observed erratic behavior of transparent powders.[17] It is now known that impurities of relatively high absorption coefficient for the radiation transmitted through transparent powders are responsible for local overheating which causes ignition and erratic performance of the propellant.

Calculations of heat transfer by radiation to a burning propellant require not only knowledge of the radiant energy emitted from the powder gases but also information on the absorption coefficient of solid powders. The absorption coefficient K_λ is defined by the relation

$$I/I_0 = \exp(-K_\lambda x), \tag{5}$$

where I/I_0 represents the fraction of radiant energy in the wavelength region between λ and $\lambda + d\lambda$ which is transmitted through a thickness x of the powder. Extensive experimental determinations of K_λ have been carried out by Daniels *et al.*[18] and by Crawford *et al.*[19] As the result of this work, reasonably accurate values for K_λ are available for transparent double-base powders and also for powders containing finely divided carbon added as darkening agent.

Various treatments have been developed for estimating the effect of radiation on the performance of solid-fuel rockets by calculating the rise in powder temperature associated with the absorption of radiant energy and by determining the corresponding change in the rate of burning. Beek[11] considered the refraction of radiation through a homogeneous propellant and included in his treatment the radiation incident from all directions. Crawford and Parr[20] similarly considered the radiant heat transfer to a small cylindrical grain burning in an inert atmosphere. The treatments of Beek and of Crawford and Parr involve geometric and computational difficulties which can be avoided, to some extent, by including only the radiant energy incident in a direction normal to the powder surface. This procedure is justified, as a first approximation, if the absorption coefficient of the powder is suffi-

[17] W. H. Avery, Paper 24 presented before the *Symposium on kinetics of explosives* at the 112th meeting of the American Chemical Society in New York City, September 1947.

[18] F. Daniels and S. S. Penner, OSRD Report No. 6559, December 1945.

[19] B. L. Crawford, Jr. and R. Gledhill, OSRD Report No. 6374, December 1945.

[20] B. L. Crawford, Jr. and R. G. Parr, University of Minnesota, Report No. 27, February 1945. This report is available in the files of Division 3, Section H, NDRC.

ciently high, since in this case radiant energy passing through the powder at nonnormal incidence is attenuated so rapidly that it will contribute very little to the radiant energy absorbed at some distance from the initial powder surface. Avery[21] and Dresher and McClure[22] were the first to introduce this simplification. Their treatments are fundamentally alike. Avery used a constant burning rate and a variable pressure, whereas Dresher and McClure carried through their analysis by assuming a constant pressure but a variable burning rate. An iterated procedure for using a burning-rate constant over short distances and a variable pressure was later introduced.[23]

Application of the theories for the effect of radiation on the performance of solid-fuel rockets has been made to a number of problems. Beek[11] calculated the temperature rise in perforated powder grains of service rockets, whereas Crawford and Parr[20] showed that the heat transfer by radiation to powder strands burning in an inert atmosphere in an apparatus for measurement of the burning rate of solid propellants[24] was negligibly small. Avery[21] correlated the chamber pressure of a solid-fuel rocket motor with the time of burning and developed a simple method for calculating the rise in powder temperature associated with the absorption of radiant energy. Dresher and McClure[22] deduced a procedure for determining the burning rate as a function of the fraction of powder already burned and also showed how to calculate the temperature rise associated with the absorption of radiant energy. Avery's method was later employed for estimating an effective mass emissivity from the nature of the pressure *vs.* time curve in some transparent powders.[23] The result was found to be in reasonably good agreement with Craig's experimental values. Extensive calculations have also been made to show that the theoretical increase in burning rate, which is associated with the increased absorption of radiant energy caused

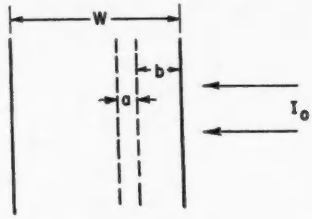

FIG. 1. Heating of a powder slab of width W by radiant energy incident from one side in a direction normal to the surface of the powder.

by replacement of a small test motor with a larger unit, is of the order of a few percent. This conclusion is in agreement with experimentally observed results.[23]

The theoretical treatments have shown that the effect of radiation on the performance of solid-fuel rockets can be used to account for a number of observed experimental facts. An attempt to make a direct experimental check for the dependence of average burning rate on average radiation path length, by using step-machined powder grains,[23] gave qualitative confirmation of the radiation theory. Repetition of this work with the use of much larger variations in path length at lower chamber pressures, in order to magnify the effect of radiation, appears desirable.

Only one application of the radiation theories to liquid-fuel rockets appears to have been made so far.[25] Since there is considerable doubt regarding the nature of the radiant energy emitted by the gases in liquid-fuel rockets, the theories which have been shown to be applicable to solid-fuel motors should be used for liquid-fuel rockets with considerable reserve. The geometrical problem solved for the radiant heat transfer to a liquid droplet moving along the axis of a cylindrical motor chamber[25] can be readily adapted to a nongraybody radiating environment.

In concluding this survey of the application of the theory of the effect of radiation on rocket performance, it appears desirable to illustrate the technique used for calculating the temperature rise associated with the absorption of radiant energy by a simple example taken from the work of Avery.[21]

Consider a layer of powder parallel to the

[21] W. H. Avery, OSRD Report No. 3880, July 1944; Part II, OSRD Report No. 5817, 1946.

[22] M. J. Dresher and F. T. McClure, Internal Memorandum of Division 3, Section H, NDRC, July 1944; Part III, OSRD Report No. 5817, 1946.

[23] S. S. Penner, OSRD Report No. 5251, 1945; Part IV, OSRD Report No. 5817, 1946. *J. App. Physics* **19**, 278 (1948); *ibid*, **19**, 392 (1948); *ibid*, **19**, 511 (1948).

[24] B. L. Crawford, Jr., and C. Huggett, OSRD Report No. 6374, 1945; F. Daniels and R. E. Wilfong, OSRD Report No. 6559, 1945.

[25] S. S. Penner and S. Weinbaum, *J. Opt. Soc. Am.* **38**, 599 (1948); *ibid.*, **38**, 840 (1948).

burning surface, of unit cross section and thickness a at a distance b from the initial powder surface (Fig. 1). Radiant energy of intensity, I_0 is impinging in a direction normal to the powder surface from one side. It is desired to calculate the temperature rise of unit cross-sectional area of width a during the time required for the powder to burn through the width b. If X represents the distance of the layer a from the burning powder surface and r is the constant rate of burning of the powder in a direction normal to its surface, then

$$X=b-rt, \tag{6}$$

where t represents the time. The energy dE_a absorbed in a during the time dt is

$$\begin{aligned} dE_a &= \{I_0 \exp(-KX) - I_0 \exp[-K(X+a)]\}dt \\ &= I_0[1-\exp(-Ka)] \exp(-KX)dt, \end{aligned}$$

where K is the absorption coefficient of the powder in cm^{-1}, which is assumed to be independent of wavelength in the present case. The total radiant energy E_a absorbed in time t is

$$E_a = I_0[1-\exp(-Ka)] \int_0^t \exp(-KX)dt. \tag{7}$$

The limits of integration are determined by noting that

$$X=0 \text{ when } t=b/r,$$
$$X=b \text{ when } t=0.$$

Hence

$$\begin{aligned} E_a &= I_0[1-\exp(-Ka)] \\ &\quad \times \exp(-Kb) \int_0^{b/r} \exp(Krt)dt \\ &= I_0[1-\exp(-Ka)][1-\exp(-Kb)]/Kr. \end{aligned} \tag{8}$$

The temperature rise ΔT of the powder is

$$\Delta T = E_a/Ca\rho, \tag{9}$$

where C is the heat capacity of the powder and ρ is the powder density. For sufficiently small values of a,

$$[1-\exp(-Ka)] \cong Ka,$$

and therefore

$$\Delta T = I_0[1-\exp(-Kb)]/rC\rho.$$

The calculated increase in powder temperature, associated with the absorption of radiant energy, can be related to the rate of burning of the propellant. In this manner, changes in experimentally observable performance characteristics can be calculated from the radiant energy absorbed by a propellant grain.

The preceding calculations have been extended to the case where I_0 varies with time because the radiation path length is increasing and where K depends on the wavelength.[11, 21–23]

Conclusions

The subject of the effect of radiation on the performance of solid-fuel rockets appears to have been discussed adequately as far as double-base propellants are concerned. Further refinements of the theory can hardly be expected to yield important useful information, since the influence of radiation on burning rates is of secondary importance. The important practical problem of adequate darkening for transparent powders was solved several years ago through the work of Avery, Beek, McClure, Daniels, Crawford, and others.

No experimental information regarding the nature of the radiation from the flames of composite propellants[1] appears to be available. On the basis of the results obtained by Kracek and Benedict it might be expected that the relatively high concentration of solid particles in the flame will favor graybody or blackbody emission characteristics. The effective absorption coefficient for radiant energy of composite propellants should be very high because of multiple scattering through a heterogeneous powder matrix. It is probably for this reason that problems such as radiation fissuring have not arisen in connection with the development of composite propellants.

As was mentioned earlier in this discussion, relatively little is known about the emission of radiant energy from liquid-fuel motors. It is therefore difficult to estimate radiant heat transfer to the motor chamber or to liquid droplets. However, considerable information regarding the emission characteristics of various propellant combinations should be forthcoming as the result of combustion studies now under way at various laboratories.[13, 14] Other work of interest in connection with radiation studies from rockets is concerned with the measurement of infra-red ab-

sorption coefficients of gases at various pressures.[26] It is planned at the Jet Propulsion Laboratory to extend the studies of absorption coefficients in the infra-red region to pressures of 20 atmospheres, corresponding to the normal operating pressure of liquid-fuel rocket motors. The results of this work should be of interest also in connection with fundamental investigations of line broadening caused by foreign perturbers.[6]

[26] A. M. Thorndike, *J. Chem. Physics* **16**, 211 (1948).

Education of Physicists for Petroleum Exploration and Production

D. H. Clewell
Magnolia Petroleum Company, Dallas, Texas

WHEN one writes about training students for geophysical work, it is often true that two different kinds of geophysicists are being considered. One type has on occasion been labeled the *geophysical engineer*, meaning a person engaged in the straightforward problem of searching for oil using the known tools of the trade, such as the seismograph, gravity meter and magnetometer. The other type might be known as a *geophysicist* and be described as a person more interested in research on new and improved methods and on the development of new principles of interpretation. The distinction is the same as that between a chemist and chemical engineer or as between a physicist and an electrical or mechanical engineer.

In the case of the geophysical engineer and the geophysicist I am of the strong opinion that it is unnecessary to draw much distinction between the two, particularly in the *training of prospective employees*. It cannot be denied that the actual work of finding oil will naturally divide itself into these two categories, and, after a man has had a number of years of experience in the industry, he will find his niche either as a geophysical engineer or a geophysicist.

However, a discussion of the type of training desired by various supervisors in a typical oil company revealed the fact that all of them hoped for essentially the same kind of training in the young college graduate, whether they were searching for laboratory men or for staff to man field parties. Minor differences in the training preferred came to light, but I doubt that these differences are more than would normally be expected in the variation of one student from another in the selection of elective subjects or thesis topics. Perhaps the main difference in the training required of the so-called geophysical engineer and the geophysicist is the fact that the latter will, as a rule, profit more from advanced study leading to the master's or doctor's degree than the man who contemplates becoming a geophysical engineer. Again the difference is not so much in the type of training as in the amount.

In the first place, a sound training in classical physics is of the utmost importance. Probing into the earth with sound waves to determine subsurface structures, or attempting to infer these structures from the way they distort the earth's gravitational or magnetic fields, demands familiarity with the fundamentals of wave propagation and of potential theory. A good solid course in theoretical physics, as given in Page's[1] or in Slater and Frank's textbooks,[2] is very desirable. It is, of course, difficult to find time to bring students up-to-date on all the modern developments in physics without slighting some of the older courses. Training in modern physics is essential but, nevertheless, in the preparation for a career in geophysics the primary emphasis must remain on classical physics. Other highly desirable courses are acoustics, mechanics, optics and photography, electricity and magnetism, atomic structure and descriptive geometry. To this list should be added inorganic, organic and analytical chemistry.

[1] L. Page, *Introduction to theoretical physics* (Van Nostrand, 1935).

[2] Slater and Frank, *Introduction to theoretical physics* (McGraw Hill, 1933).

Mathematics is obviously essential to any constructive work in physics or geophysics. The minimum desirable course of study should include differential equations. As a matter of fact, a physicist with a minor in mathematics is at present a much sought for man for many phases of geophysical work.

Since the geophysicist deals with the earth, it is self-evident that a physicist planning to enter the geophysical industry can derive great benefit from courses in geology. Such courses should, if possible, be extensive enough to include structural geology and sedimentology. A minor in geology is particularly desirable for field work in geophysics.

Other courses of special importance exist in the field of electronics. Courses on communication circuits, and vacuum-tube applications, frequently given in an electrical engineering department, are very helpful to those who prefer to deal mostly with geophysical prospecting tools and equipment. A physicist with a minor in communications is well fitted for laboratory and maintenance work.

Finally, but far from last in importance, is English composition. There is no doubt that many good technical men fail to be promoted simply because they cannot express themselves well and make themselves understood clearly. If a man cannot express his thoughts clearly to others, there is always a possibility of doubt concerning his own clear understanding of a problem.

So far it should be noted that there has been no mention of courses in geophysics itself. This omission is intentional, and emphasizes the fact that a strong fundamental training in physics, chemistry and mathematics is very desirable and essential. The fundamental sciences providing the background to geophysics cannot be learned easily after leaving college, whereas current practices in geophysics can be learned even better in industry than in school. This attitude is, of course, not new, but it deserves repetition. However, I am certain that men in the petroleum industry would still recommend that a short course in geophysics be offered to the undergraduate for the express purpose of acquainting him with geophysics as a field of endeavor in which his broad training in fundamentals has many varied and interesting applications. Although it is not so true of the Southwestern schools, we have found that in many Eastern and Midwestern colleges the students, especially physicists and electrical engineers, have never realized that opportunities exist for them in geophysics. A course to whet their appetite for this work would be very beneficial in attracting more students into the field.

It is obviously impossible to cover thoroughly in a four-year course all of this rather long list of subjects, however desirable they may be for students entering the geophysical industry. My own belief is that the material could be distributed better over a five-year course. Such a situation is not peculiar to geophysical training. The amount of information that must be brought to the attention of the technical student has been increasing for years, and at present is increasing at a faster tempo than ever. Just as industry in the past has placed a greater premium on four-year college training, I believe that, in the very near future, the five-year course will tend to become the acceptable minimum. By five-year course, I am referring to a course somewhat similar to the normal requirements for a master's degree, but differing therefrom in that the five-year curriculum be an integrated and interrelated program of subjects.

As was mentioned above, the requirements for a man entering *research* work in geophysics are much the same as those already suggested except that the training should preferably reach a more advanced stage and perhaps lead to a doctor's degree. In the advanced graduate training it is believed that a number of specialized courses in geophysics and geology could be included. A doctor's thesis on some geophysical problem would be quite appropriate.

The discussion, so far, has been pointed at the training of men for geophysical work. However, the suggested training will fit a man equally well for many aspects of *production research* in the oil industry, for example, improving methods of extracting oil and gas from the ground. This is especially apparent in the field of reservoir behavior where one deals with the flow of fluids through porous media. The basic mathematical principles of fluid flow are very similar to those of gravitational and magnetic fields and of the

propagation of energy by wave motion. Also there is no fundamental difference between the designer of geophysical instruments and the designer of many of the instruments required by the petroleum engineer. A man interested in electronics can find interesting problems in either geophysical or production research.

In conclusion, I should like to raise two rather general points. The first is concerned with the belief that more attention should be given by educators to the encouragement of a *creative* approach on the part of the new graduates. At the present time, it would seem that there is a heavy concentration on *analysis*. Students are taught to break a problem or an observed phenomenon into its component parts. In so doing it is demonstrated that many different phenomena are based on the same few fundamental principles. By this procedure of analyzing observations into their fundamental components we are enabled to understand many more facts than we could otherwise. The value of such analytical ability is unquestioned.

However, progress in any science or industry must also rely on a second step—namely, the *synthesis* of fundamental components into new and useful ideas or machines. The synthesis is just as important as the analysis, and yet it is my experience that the creative individual who can visualize new combinations of old and known ideas is much rarer than the skilled analyst who can always take a problem apart and tell where the difficulty or contradiction lies.

Is it possible to introduce into our educational system a method for evaluating this creative factor? I doubt that we now know how to teach a man to be creative, but it should be possible to recognize creative ability in the sciences, and then encourage those who possess it to develop it to the fullest.

The second general point is somewhat related to the first. It is simply an expression of an employer's desire to know, prior to the hiring of a young graduate, something about his ability to carry on independent work. Can he think for himself? Is he the type of man who has the proper blend of knowledge and perseverance to bring his work to some definite conclusion in a reasonable length of time? I wish the solution to this problem were known. The only suggestion I can make is that a student should be tested on an original thesis problem. My personal opinion is that more can be learned about a prospective employee by discussing his thesis with him than in any other way. The manner in which an applicant presents the objectives of his thesis, the detailed problems that arose and his course of action in solving these problems give an excellent insight into his potentialities as a member of our staff.

Automatic Weather Station

Making the weather do the work of reporting itself by harnessing the wind is the goal of an experiment in progress at the Signal Corps Engineering Laboratories at Fort Monmouth, New Jersey. Having developed an automatic weather station that can send radio reports on temperature, pressure, relative humidity, wind speed, wind direction, precipitation and sunshine intensity, the Signal Corps now is seeking to extend its period of unattended operation beyond one year. If current experiments are successful, the weather stations will be placed in near-inaccessible corners of the world and allowed to run themselves.

The problem has been attacked by combining modern electronics with the centuries-old windmill, which charges a bank of storage batteries. The batteries in turn operate the automatic weather station. A wind of 7 miles per hour is sufficient to generate electricity, and one of 24 miles per hour will produce the generator's rated output of 2.5 kilowatts. An automatic regulator will prevent over-charging of batteries.

Instead of the picturesque arms usually associated with windmills, the Signal Corps currently is trying a three-bladed propeller. The height of the tower supporting the generator will vary in accordance with its location and the intensity of winds there. The equipment has already undergone preliminary tests on Mt. Washington, New Hampshire.

Cold and ice present problems that are being investigated. Tests have been made of de-icing equipment, and consideration is being given to placing the batteries in an underground vault to keep them from freezing. They may go as deep as 20 feet below the earth's surface where the temperature remains more or less constant. The need for frequent refilling has been eliminated by furnishing batteries of larger volume.

Some Consequences of a Simple Theorem on Torque

HOWARD EVES
Oregon State College, Corvallis, Oregon

THE following paper develops a simple theorem on torque and derives from this basic one a number of additional physical and geometrical theorems and constructions. Consideration is paid to the diametral line of a system of lines for a given direction and a given system of multiples. The concept of diametral line is applied to the problem of locating the point with respect to which a coplanar nonparallel system of forces possesses minimum inertia. This, in turn, leads to some new theorems in modern geometry.

The Theorem

In the sequel it will be found convenient to speak of the positive and negative sides of a directed line lying in a given plane. The positive side of such a line is that portion of the plane containing points about which the directed line exerts a positive, or counterclockwise, rotation, when the plane is viewed from a given side. The other side of the line will be known as the negative side. A perpendicular dropped from a point in the given plane upon a directed line will be considered as positive or negative according as the point lies on the positive or negative side of the line.

We shall be concerned with the torques of a number of coplanar forces about an axis perpendicular to the plane of the forces. Since no misunderstanding can arise, we shall, for convenience, speak of the torques of the forces about a point P, meaning, of course, the torques of the forces about an axis through P and perpendicular to the plane of the forces. The following theorem constitutes the foundation of this note.

Let F_i $(i=1, \cdots, n)$ be a set of n coplanar forces not in equilibrium. Then the locus of a point P in the plane of the forces, such that the sum of the torques of the forces about P is a constant k, is a straight line. As we vary k the line moves parallel to itself.

We can establish this theorem very readily by purely physical reasoning (as will be indicated later in *corollary D*), or by a geometrical argument. For future purposes we here choose the latter procedure, and to this end select some rectangular Cartesian frame of reference in the plane such that the origin of coordinates does not lie on the line of action of any one of the forces. Let the equation of the line of action of the force F_i be normalized to give

$$a_ix+b_iy+c_i=0, \quad a_i^2+b_i^2=1, \quad c_i>0.$$

Further, suppose each line to be directed by the direction of the force on the line. Let f_i be the (positive) magnitude of the force F_i and let (u, v) be the coordinates of the point P. Then the torque of the force F_i about point P is given by

$$f_ie_i(a_iu+b_iv+c_i),$$

where $e_i=\pm 1$ according as the origin is or is not on the positive side of the line. We require that

$$\sum f_ie_i(a_iu+b_iv+c_i)=k,$$

and thus find that the locus of P is the straight line

$$(\sum f_ie_ia_i)x+(\sum f_ie_ib_i)y+(\sum f_ie_ic_i)-k=0. \quad (1)$$

Since $\sum f_ie_ib_i$ and $-\sum f_ie_ia_i$ represent the sums of the components of the forces on the x and y axes, respectively, and since the forces are assumed to be not in equilibrium, it follows that $\sum f_ie_ia_i$ and $\sum f_ie_ib_i$ cannot vanish simultaneously. Thus the straight line Eq. (1) is a line in the finite part of the plane, and by varying k we obtain a family of parallel lines having slope

$$m=-(\sum f_ie_ia_i)/(\sum f_ie_ib_i).$$

Some Consequences

As immediate corollaries to the preceding theorem we have the following:

Corollary A. The locus of a point P about which an unbalanced coplanar system of forces has zero torque is a straight line, namely, the line obtained by setting $k=0$ in Eq. (1) above.

We shall call the locus of *corollary A* the *axis* of the unbalanced coplanar system of forces F_i.

Corollary B. Given two unbalanced systems of forces in the same plane, then, in general, there

exists a unique point P in the plane about which each system exerts a given torque, namely, the point of intersection of the appropriate straight line loci of P for the two unbalanced systems of forces.

Corollary C. Given an unbalanced coplanar system of forces, then there exists a line of points P in the plane about which a given part of the system exerts a torque equal to the negative of that exerted by the remaining part of the system, namely, the axis of the combined system of forces.

Corollary D. The line of action of the resultant of an unbalanced coplanar system of forces coincides with the axis of the system of forces.

For, by *Varignon's theorem*, the sum of the torques of the forces about a point P in their plane is equal to the torque of the resultant of the forces about P. But the locus of points P about which the resultant has zero torque is obviously the line of action of the resultant, and the locus of points P about which the system has zero torque is the axis of the system. Therefore the line of action of the resultant coincides with the axis of the system.

As purely geometrical theorems we have the next three corollaries.

Corollary E. The locus of a point P for which the sum of given multiples of the signed perpendiculars from P upon a system of coplanar directed lines is a constant k, is, in general, a straight line. As k varies, the line moves parallel to itself.

When the constant vanishes we may define the resulting line as the *axis* of the system of coplanar directed lines for the given system of multiples. We then have the following construction for the axis of a system of coplanar directed lines for a given system of multiples.

Corollary F. Lay off a vector on each line, having its length proportional to the corresponding multiple, the direction of the vector being the same as or opposite to that of the line according as the multiple is positive or negative. The line of action of the resultant of these vectors is the required axis.

Corollary G. The locus of a point P for which the sum of the signed areas subtended at P by a number of coplanar directed line segments is a constant k, is, in general, a straight line. As k varies, the line moves parallel to itself.

Take the lengths of the line segments as a system of multiples, and apply *corollary E.*

Corollary H. Let us be given a system of n coplanar lines L_i cut by a variable line L in the points A_i. Consider masses m_i located at the points A_i. Then, as L moves parallel to itself, the locus of the center of gravity P of the masses m_i is a straight line.

Direct the lines L_i and L in any manner. Let θ_i be the counterclockwise angle from the positive end of L to the positive end of L_i, and let a_i be the signed perpendicular distance of P from L_i. Then

$$PA_i = a_i \csc\theta_i.$$

But, by hypothesis,

$$\sum m_i PA_i = 0.$$

Hence we find that

$$\sum (m_i \csc\theta_i) a_i = 0,$$

whence P describes, by *corollary E*, the axis of the lines L_i for the system of multiples $m_i \csc\theta_i$.

The locus of *corollary H* will be called the *diametral line* of the system of lines L_i for L and the system of multiples m_i.

A very interesting problem is the determination of the envelope of all diametral lines of a given system of lines for a given system of multiples. The problem for the general case is not easy. Let it suffice to state here without proof that the diametral lines of the sides of an equilateral triangle for a system of equal multiples envelop the incircle of the triangle.

Inertia with Respect to a Point

We shall close this paper with an application of the foregoing to a problem on inertia. We define the moment of inertia of a force F with respect to an axis perpendicular to F as the product of the magnitude of F by the square of the perpendicular distance between F and the axis. We shall consider the moments of inertia of a number of coplanar forces with respect to an axis perpendicular to the plane of the forces. For brevity we shall speak merely of the inertia of the forces with respect to a point P, meaning, of course, the moments of inertia of the forces

with respect to an axis through P and perpendicular to the plane of the forces. We now prove the following theorem:

Let F_i $(i=1, \cdots, n)$ be a set of n coplanar nonparallel forces. Then the locus of a point P in the plane of the forces such that the sum of the inertias of the forces with respect to P is a constant k, is an ellipse. As we vary k we obtain a family of similar ellipses having a common center and common orientation.

Let us normalize and direct the lines of action of the forces F_i as at the beginning of the paper, and let (u, v) denote the coordinates of P. Then we require that

$$\sum f_i[e_i(a_iu+b_iv+c_i)]^2=k.$$

The locus of P is thus the conic

$$\begin{aligned} G(x, y, k) \equiv (\textstyle\sum f_ia_i^2)x^2+2(\sum f_ia_ib_i)xy \\ +(\textstyle\sum f_ib_i^2)y^2+2(\sum f_ia_ic_i)x \\ +2(\textstyle\sum f_ib_ic_i)y+\sum f_ic_i^2-k=0. \end{aligned}$$

Since the forces are nonparallel, the a_i's are not proportional to the b_i's, and Cauchy's (Schwarz's) inequality guarantees that

$$\sum f_ia_i^2 \sum f_ib_i^2-(\sum f_ia_ib_i)^2>0.$$

We thus see that the conic is an ellipse. Also, since the quadratic terms are independent of k, by varying k we obtain a family of similar and similarly situated ellipses. Finally, we observe that the center of the conic $G(x, y, k)=0$ is given as the intersection of the lines $\partial G/\partial x=0$ and $\partial G/\partial y=0$. Since these partial derivatives are free of k it follows that all the ellipses have a common center. The theorem is now established.

The common center of the ellipses is obviously the point P with respect to which the system of forces possesses minimum inertia. The concept of diametral line, developed in the past section, furnishes us with a construction locating this point P. We therefore consider the problem:

Find the point P with respect to which a given system of coplanar nonparallel forces possesses minimum inertia.

Let L_i $(i=1, \cdots, n)$ be the lines of action of the forces and f_i $(i=1, \cdots, n)$ their (positive) magnitudes. Let L be any line through P meeting the lines L_i in the points A_i at the angles θ_i. Let p_i denote the perpendicular distance of P from L_i. Then

$$\sum f_ip_i^2=\sum f_i \sin^2\theta_i(PA_i)^2.$$

This will be a minimum when

$$\sum f_i \sin^2\theta_i PA_i=0,$$

or, in other words, when P is the center of gravity of masses $f_i \sin^2\theta_i$ placed at the points A_i. Let L' be the diametral line of the system of lines L_i for L and the system of multiples $f_i \sin^2\theta_i$. Then L' passes through the required point P. Let L' meet the lines L_i in the points A_i' at angles θ_i'. Then P is the center of gravity of masses $f_i \sin^2\theta_i'$ placed at the points A_i'.

It has been known for a long time that if we have only three forces F_i, of equal magnitudes, then the point P of the problem is the symmedian point of the triangle determined by the three lines L_i.[1] The problem thus involves a generalization of the symmedian point of a triangle to an analogous point for a plane polygon of any number of sides. We can also now state some theorems about the symmedian point of a triangle which do not occur in the ready literature; for example:

Let L, any line drawn through the symmedian point P of a triangle ABC, cut the sides BC, CA, AB in points A', B', C' at angles α, β, γ. Consider particles of masses $\sin^2\alpha$, $\sin^2\beta$, $\sin^2\gamma$ placed at A', B', C'. Then P is the center of gravity of these masses.

[1] R. A. Johnson, *Modern geometry* (Houghton Mifflin, 1929), p. 216.

One may also imagine, that in criminal hands radium might become very dangerous, and here we may ask ourselves if humanity has anything to gain by learning the secrets of nature, if it is ripe enough to profit by them, or if this knowledge is not harmful. The example of Nobel's discoveries is characteristic: powerful explosives have permitted men to perform admirable work. They are also a terrible means of destruction in the hands of the great criminals who lead the peoples towards war.

I am among those who think, with Nobel, that humanity will obtain more good than evil from the new discoveries.—PIERRE CURIE, 1905.

AMERICAN JOURNAL OF PHYSICS VOLUME 16, NUMBER 9 DECEMBER, 1948

NOTES AND DISCUSSION

A New Type of Viscosimeter

A. A. ELKARIM
Farouk I University, Alexandria, Egypt

DESCRIPTION of the Apparatus.—The essential features of the new type of viscosimeter may be seen in Fig. 1. It consists of a soft glass capillary tube with two marks, *A* and *B*. The tube is connected to two glass cylinders, one at each end, provided with stopcocks. The bends in the tube serve to impede the escape of the air bubble when the pressure in one limb exceeds that in the other.

Before being filled, the apparatus is thoroughly cleaned with chromic acid followed by steam, and is finally dried with hot filtered air. It is supported by a wooden holder clamped to the bench, adjusted so as to be dead level. The mean radius of the capillary tube is ascertained from the weight of the mercury needed to fill it. The liquid, whose coefficient of viscosity we wish to measure, is introduced into the capillary tube with an air bubble as index. In order to drive the index from end to end, the stopcocks T_1 and T_2 are opened simultaneously. The terminal velocity of the index corresponding to any particular difference in pressure between the ends is obtained by timing its passage with a stop watch, the timing being for the interval from the instant the left-hand end of the bubble is at the mark *A* to the instant when the right-hand end is at mark *B*. Velocities greater than 15 cm per sec could not be measured sufficiently accurately owing to the difficulties in timing, while at very low velocities the liquids adhered. The graph Fig. 2 shows those velocities obtained within the range where accurate observation is possible.

Theoretical Details.—The flow of liquids through capillary tubes has been the subject of many investigations, the main problem being based on a law laid down by Jean Marie Poiseuille, whose work was published in 1842. Poiseuille's equation for an incompressible fluid may be represented in the formula

$$P=(8\eta LV/\pi R^4), \tag{1}$$

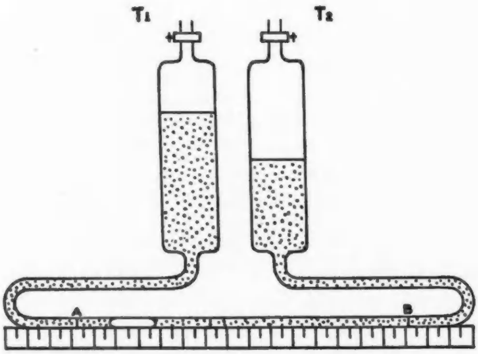

FIG. 1. A new type of viscosimeter.

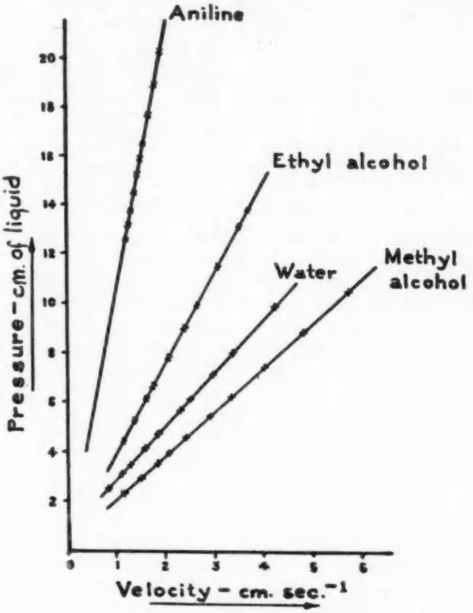

FIG. 2. The variation of pressure difference with velocity for various liquids.

where P is the difference of pressure at the two ends of the tube in dyne cm^{-2}, η is the coefficient of viscosity in dyne sec cm^{-2}, L is the length of the tube in cm, V is the rate of flow in cm^3 sec^{-1}, and R is the radius of the tube in cm. Take an air index at rest in a horizontal tube, sufficiently narrow for its end surface to be considered spherical. Let the pressure in the left-hand cylinder be greater than that in the right-hand cylinder by an amount P_1. Therefore,

$$P_1=h_0\rho g,$$

h_0 being the difference in height, ρ the density of the liquid, and g the acceleration due to gravity. The index therefore tends to move from left to right with uniform velocity. The total pressure difference between the ends of the tube when the index is in motion is seen to embody three parts—a pressure difference due to (a) the viscous forces brought into play by the relative motion of the different layers of the liquid, (b) the viscous forces stimulated by the relative motion of the different layers of the air index, (c) the difference in curvature between the two ends of the air index. If this pressure difference be denoted by ϵ, then

$$\epsilon=\frac{2T}{R}(\cos\theta_1-\cos\theta_2),$$

T being the surface tension of the liquid, θ_1 the angle of contact at the left end, and θ_2 the angle of contact at the

TABLE I.

Liquid	ϵ (dyne/cm²)	Density ρ (g/cm³)	($\cos\theta_1 - \cos\theta_2$)	Surface tension T (dyne/cm)	Viscosity η (cgsu)	Temperature θ (°C)
Water	$0.65\rho g$	0.998	0.26	72	0.0091	25
Methyl alcohol	$0.25\rho g$	0.795	0.25	23	0.0055	25
Aniline	$0.4\rho g$	1.02	0.27	43	0.042	21
Ethyl alcohol	$0.25\rho g$	0.793	0.26	22	0.012	21

right end of the index. Thus for the mean total pressure difference between the ends of the tube we have

$$P = \{k_0 - [a(x-l)/A]\}\rho g, \tag{2}$$

a being the cross-sectional area of the capillary tube, A the cross-sectional area of the cylinder, l the length of the air index, and x the displacement of the index in the horizontal tube between A and B. Poiseuille's equation for an incompressible fluid may be represented in the formula

$$P = \frac{8\eta(L-l)v}{R^2}, \tag{3}$$

L being the length of the tube, R its radius, and v the velocity. Now since the viscosity η' of air is small in comparison with the viscosity η of the liquid, the same form of this equation may be used for the flow of air as for the flow of liquid. Thus, to maintain this flow, the fall in pressure along the tube is $(8\eta' vl/R^2)$ which remains constant during the movement of the index along the tube. Therefore, we have

$$P = (8v/R^2)[l\eta' + (L-l)\eta] + \epsilon. \tag{4}$$

If now the variation of the difference in pressure P is represented on the y axis, and the variation of velocity is represented on the x axis, the graph between P and v will be a straight line, whose intercept on the y axis will be ϵ, and from its slope the coefficient of viscosity can be determined.

Conclusions.—Experiments have been carried out on water, methyl alcohol, aniline, and ethyl alcohol at room temperature kept perceptibly constant during the half-hour needed for any one set of observations. Table I provides the final results. For all the observations a fine glass capillary tube 0.058 cm in radius and 100 cm in length was used. The graph in Fig. 2 represents the variation of the pressure difference with the velocity of air index for the respective liquids. The apparatus shown in Fig. 1 is cheap and simple to construct.

The author wishes to acknowledge the kind advice and encouragement given him by Dr. M. A. Elsherbini during these investigations.

New Members of the Association

The Following persons have been made *members* or *junior members* (*J*) of the American Association of Physics Teachers since the publication of the preceding list [*Am. J. Physics* **16**, 420 (1948)].

Bedwell, Thomas H., Florida State University, Tallahassee, Fla.
Bingaman, George C., 827 Mocksville Ave., Salisbury, N. C.
Birtley, Willard Bebb, 418 Windcrest, State College, Pa.
Box, D. S., 348 Gower Rd., Sketty, Swansea, Glam., England.
Chudom, Peter F., 4236 Broadway, Chicago 13, Ill.
Conrad, Maurice A. (*J*), 6210 Sellers St., Pittsburgh 6, Pa.
Drumheller, Carl E., 524 N. Burrowes St., State College, Pa.
Duff, James F. K., 512 Ontario St., SE, Minneapolis 14, Minn.
Gilvarry, John J., N. Am. Aviation, Inc., Los Angeles, Calif.
Harvey, David E. (*J*), 807 S. Spaulding, Chicago 24, Ill.
Hinze, Donald W. (*J*), 1782 St. Anthony Ave., St. Paul 4, Minn.
Hollingsworth, Jack R., 715 Morgan Ave., Drexel Hill, Pa.
Hopkins, Nigel J. (*J*), 1082 Argyle Ave., Verdun P. Q., Montreal, Canada.
James, Floyd E., Presbyterian Junior College, Maxton, N. C.
Jaynes, Edwin T., Graduate College, Princeton, N. J.
Kronschnabel, George James, Jr. (*J*), St. Thomas College, St. Paul 1, Minn.
Kuhlthau, Alden Robert, University of New Hampshire, Durham, N. H
Kupferberg, Kenneth M., 146–53 Delaware Ave., Flushing, N. Y.
Lawson, William L., 360 Furman St., Brooklyn 2, N. Y.
Lindstrom, Anton, 4006 Harrison St., Milwaukie, Ore.
Miller, Robert J., 328 Highland Parkway, Kenmore 17, N. Y.
Mohville, Lawrence J., John Carroll University, Cleveland 18, Ohio.
Newton, Carol M. (*J*), 2535 Regent St., Berkeley, Calif.
Patterson, William, Jr. (*J*), 526 Race Ave., Lancaster, Pa.
Porter, Warren E., 8134 Calumet Dr., Munster, Ind.
Ray, James H., 223 Congress St., Brooklyn 2, N. Y.
Rockwell, R. J., 1549 Teakwood Ave., College Hill, Cincinnati, Ohio.
Rood, Joseph Lloyd, University of California, Berkeley 4, Calif.
Ross, Nicholas V., 6804 Glenwood Ave., Youngstown, Ohio.
Scott, Franklin Robert, Indiana University, Bloomington, Ind.
Van Voorhis, S. N., University of Rochester, Rochester, N. Y.
Wellman, Richard Edward (*J*), 880 Berwin St., Akron, Ohio.
Windesheim, Karl A., University of Illinois, Urbana, Ill.

The Regional Representative of the District of Columbia Section on the AAPT Executive Committee for 1948 is E. W. THOMSON. Through an error in the reading of the Annual Report, the AAPT Secretary listed the name of T. B. BROWN.

ANNOUNCEMENTS AND NEWS

Book Review

Introduction to Color. Ralph M. Evans. Pp. 340+x, Figs. 289, 15 Plates, 7×9¼ in. John Wiley & Sons, Inc., New York, 1948. Price $6.00.

The thing that would seem to be most needed by the non-technical and technical reader alike is a grasp of color as an experience, a mode of human response in its relation to the great variety of conditions upon which it is based. It has been all too easy for the physicist to think of color as a kind of sensation quite well determined by certain characteristics of radiant energy, and either to neglect the deviations from such expectations or else to have no satisfactory way of accounting for them, once admitted. The broad field of color is cluttered with *experts*, those who know in great and intricate detail some one phase or other of the subject, but are ill-adapted to deal with the subject as a whole. Color has been dealt with not only by a variety of scientists, but also by those who consider nothing but *esthetic* effects, and whose means of obtaining these has not well been reduced to rule. To fuse the contributions of all the diverse elements would be a task very worth while. The greatest need of the greatest number of readers, whether scientist or layman, whether teacher or student, is for a treatment of the subject which will aid in developing a comprehensive view of the subject. Evans' *Introduction to Color* is a marked step in this direction.

Evans proposes to give an account of color with the full recognition that "color sprawls across three enormous subjects of physics, physiology, and psychology." He has neglected none of these fields while trying to produce an exposition simple and forthright enough to be acceptable and understandable to the greatest number of readers. He states that almost no knowledge of any part of the subject of color is presumed of the reader. While this is hardly the case, there is more than sufficient material in any one major portion of the book to make it worth while for any one interested in the subject.

The selection of material and the way that it is presented are such that the book would be good collateral reading for students in each of the three major fields already mentioned.

There are twenty-one chapters. The first sets the stage by showing the interrelations of the several disciplines underlying the subject of color. The second deals with the physical nature of light and relates it to the light-sensitive systems that exist in the eye, in photronic cells and photocells, and in photographic emulsions. The many types of light sources from the sun, black-body radiators, incandescent lamps, discharge tubes, etc., to biological sources such as the firefly are discussed in the third chapter. The general nature of illumination and shadows is dealt with in the fourth. From here on, no phase of the broad subject of color is neglected. Color, of course, is defined so as to include the achromatic as well as the chromatic, and what one *does* see rather than what one *should* see. A presentation of the structure of the eye, adaptation of the visual cells, the nature of sensitivity (visibility or luminosity) curves comes next and precedes what Evans calls the *visual* variables of color.

The various modes of perceiving color are made clear. He makes plain what is meant by aperture colors, illuminant colors, illumination colors and object colors. The latter are of two kinds—surface and volume colors. Evans also points out eleven attributes of these modes of appearance. Among these are not only the usual brightness (lightness), hue and saturation, but also sparkle, transparency, etc. Such features are not recognized in the more common books on color and how things look.

Evans also includes some discussion of the two phenomena usually called brightness- and color-constancy. These terms relate to the fact that brightness and color seen by the observer pertain to the object of which they are attributes. A white piece of paper appears white in a deep shadow when the paper is reflecting much less light per unit area than some other object seen as gray.

Throughout, Evans' presentation of color is not only helpful to those in pure science, but also to those who have specific and practical problems to solve. This is mainly true in connection with the relation of color to present-day photography.

The book contains many figures and fifteen full-page color plates illustrating some of the fundamental conditions involved in the production of color effects. Many of these apply quite directly to color photography without the book being heavily biased for color-photography enthusiasts. In no other book stripped of a "technical" mode of presentation is there to be found so much help for those who wish to begin a grounding in the subject of color as in the present volume.

Although Evans is not a psychologist, he plays the role pretty well when it is to be considered that the lingo of contemporary psychology is anything but well adapted for enlightening the beginner or other individual who has obtained all of his education in other fields. So, even though Evans has 'the mind' do a great deal of explanatory work, he can be excused. It will not be until books such as this are digested that the intelligent reading public, students or oldsters, will be ready for a more 'scientific' presentation of the psychobiological.

S. Howard Bartley
Michigan State College

Author Index to Volume 16

In this index are listed the names of authors and the titles of their articles and notes. Abstracts of papers read at meetings and of articles appearing under "Digest of Periodical Literature" in the journal are not listed; however, they are indexed in the Analytic Subject Index.

Few of us take the pains to study the origin of our cherished convictions; indeed, we have a natural repugnance to so doing. We like to continue to believe what we have been accustomed to accept as true, and the resentment aroused when doubt is cast upon any of our assumptions leads us to seek every manner of excuse for clinging to them. The result is that most of our so-called reasoning consists in finding arguments for going on believing as we already do.—James Harvey Robinson, 1863–1936.

Analytic Subject Index to Volume 16

The titles of articles are disregarded, the entries being based on analyses of the contents of the original articles. (D) designates a digest of an article or an abstract of a paper read at a meeting; (T) designates title only.

To facilitate reference to any desired subject, the index is divided into sections arranged alphabetically. The titles of these sections and of cross references to them are as follows:

It is an erroneous impression, fostered by sensational popular biography, that scientific discovery is often made by inspiration—a sort of coup de foudre—from on high. This is rarely the case. Even Archimedes' sudden inspiration in the bathtub; Newton's experience in the apple orchard; Descartes' geometrical discoveries in his bed; Darwin's flash of lucidity on reading a passage in Malthus; Kekulé's vision of the closed carbon ring which came to him on top of a London bus; and Einstein's brilliant solution of the Michelson puzzle in the patent office in Berne, were not messages out of the blue. They were the final co-ordinations, by minds of genius, of innumerable accumulated facts and impressions which lesser men could grasp only in their uncorrelated isolation, but which—by them—were seen in entirety and integrated into general principles. The scientist takes off from the manifold observations of predecessors, and shows his intelligence, if any, by his ability to discriminate between the important and the negligible, by selecting here and there the significant stepping-stones that will lead across the difficulties to new understanding. The one who places the last stone and steps across to the terra firma of accomplished discovery gets all the credit. Only the initiated know and honor those whose patient integrity and devotion to exact observation have made the last step possible.—Hans Zinsser, 1878–1940.

indsay

ments,

Hughes,

Eberle

th, and

ag and

Howard

, R. F.

-121(D)

rburton

Spencer

25(T)

y, E. C.

e, B. H.

ide rule

74